



UNIVERSITAT^{DE}
BARCELONA

Trabajo final de grado

GRADO DE MATEMÁTICAS

Facultad de Matemáticas e Informática
Universidad de Barcelona

MODELOS ARCH Y GARCH: Aplicación a series financieras

Autor: Kevin Amate Vicente

Director: Dr. Josep Vives

Realitzat a: Departamento de Matemáticas e Informática

Barcelona, 27 de junio de 2018

Índice

1. Abstract	1
2. Introducción	2
3. Objetivos del proyecto	3
4. Consideraciones y nociones previas	4
4.1. Análisis clásico (Determinista)	5
4.2. Procesos estocásticos	6
4.3. Procesos de segundo orden	6
4.4. Procesos estacionarios	7
4.5. Ejemplos de procesos estocásticos	8
4.5.1. Ruido IID	8
4.5.2. Ruido blanco	8
4.5.3. Ruido blanco Gaussiano	8
5. Modelos ARCH y GARCH	9
5.1. Introducción a los modelos ARCH	9
5.2. Stylized facts	10
5.3. Modelo ARCH(1): Definición y propiedades	12
5.3.1. Propiedades del modelo ARCH(1)	12
5.3.2. Estimación de los modelos ARCH(1)	17
5.4. Modelo ARCH(q): Definición y propiedades	20
5.4.1. Propiedades del modelo ARCH(q)	21
5.4.2. Estimación de los modelos ARCH(q)	23
5.5. ARCH generalizado: GARCH	24
5.6. GARCH(p,q): Definición y propiedades	24
5.6.1. Propiedades del modelo GARCH(p,q)	25
5.6.2. Estimación de los modelos GARCH(p,q)	27
5.7. Extensiones de los modelos GARCH	29
5.7.1. Modelo GARCH Exponencial	29
5.7.2. Modelo Threshold ARCH	31
6. Aplicaciones de los modelos ARCH y GARCH	33
6.1. Precio del oro (Dólar/Onza)	33
6.2. Divisa Euro-GBP	38
6.3. Cotización en la bolsa de BBVA	42
6.4. Gasolina Sin Plomo (Dólar/Galón)	46
7. Conclusiones	50

1. Abstract

En este trabajo se explica la teoría relacionada con los modelos con heterocedasticidad condicional autorregresiva ARCH y GARCH, los cuales como su nombre indica se basan en el modelado a partir de la premisa de tener una variabilidad condicional que depende de valores pasados. Posteriormente se aplican técnicas para ajustar estos modelos a diversas series financieras, las más adecuadas para los modelos suelen ser los más adecuados.

In this paper we explain the theory related to the models with conditional autoregressive heterocedasticity ARCH and GARCH, which as its name indicates are based on modeling with the premise of having a conditional variability that depends on past values. Subsequently, techniques are applied to adjust these models to various financial series, the most appropriate for these models.

2. Introducción

En el último curso de mi carrera, debido al hecho de escoger la mención de estadística, he cursado la asignatura de Análisis de Series Temporales, impartida por Antoni Messeguer Arto-la. Esta ha sido para mi una de las más interesantes, y ha sido por ese motivo que he decidido enfocar mi Trabajo Final de Grado a ampliar los conocimientos obtenidos en la asignatura comentada.

Las series temporales que más me llaman la atención por su impacto en la sociedad, ya que en parte son las que tienen mayor público, son las que aparecen en el mercado financiero, debido a que estas controlan la economía mundial y a la gran cantidad de gente que invierte en estas con tan de aumentar su capital.

Sin embargo, estas series no pueden ser modeladas mediante los procesos lineales clásicos ARMA estudiados en el curso, ya que estos no tienen en cuenta una posible volatilidad condicionada a valores anteriores, un factor que tiene un papel importante para el modelado de estas series. Es por esto que en este trabajo hago un estudio de los nombrados *modelos con heterocedasticidad condicional autorregresiva* (ARCH) o la generalización de estos (GARCH), los cuales tienen en cuenta el factor mencionado anteriormente y son los idóneos para el modelado de este tipo de series.

Primero se comentan una serie de conceptos previos, extraídos del *pdf sin publicar del Máster en Big Data impartido por Josep Vives*, los cuales son necesarios para la comprensión de la teoría de este tipo de modelos.

Seguidamente se explica toda la teoría que estos tienen detrás, con la ayuda de los libros *Statics of Financial Markets* de Jürgen Franke y *ARCH Models and Financial Applications* de Christian Gouriéroux, explicando los rasgos que los caracterizan, así como los diferentes tipos que nos podemos encontrar, sus diversas definiciones y propiedades.

Por último, se aplican diversas técnicas para intentar ajustar varias series temporales financieras, utilizando los libros *Computacional Finance* de Argimiro Arratia, *Time Series Analysis and Its Applications* de Robert H. Shumway et al., *Mastering R for Quantitative Finance* de Edina Berlinger et al. y *Introduction to R for Quantitative Finance* de Gergely Daróczi, viendo primero si presentan las características propias de estos modelos, y posteriormente probando cual de estos es el que mejor se ajusta.

3. Objetivos del proyecto

- Entender la teoría de los modelos con heterocedasticidad condicional autorregresiva, es decir, aprender sus diversas definiciones, sus rasgos característicos y las diversas propiedades que tienen, así como los diferentes tipos y extensiones que hay de estos.
- Aprender qué tipo de series se ajustan a estos modelos, las transformaciones que hay que hacerles a estas para su posterior modelado y las técnicas necesarias para poder llevarlo a cabo.
- Aplicar estas técnicas a diferentes series financieras con la finalidad de modelarlas correctamente con el modelo que mejor se ajuste.

4. Consideraciones y nociones previas

La evolución del precio de la gasolina, cómo va cambiando la temperatura en un determinado sitio, la variabilidad del nivel de un lago... En la vida, nos encontramos una infinitud de datos, en los cuales el tiempo tiene un papel fundamental.

A cualquier colección de datos, indexados por el tiempo, es decir, ordenados cronológicamente, se les llama serie temporal. Suelen ser cantidades recogidas al largo del tiempo, en intervalos de la misma longitud, sobre algún determinado fenómeno. Éstos pueden ser anuales, mensuales, semanales, diarios, hora a hora, etc. Nosotros escribiremos una serie temporal como:

$$\{x_i, i = 0, 1, \dots, n, \dots\} \quad (4.1)$$

La idea principal para analizar este tipo de datos es la dependencia estocástica, ya que el orden juega un papel importante, haciendo que x_1, \dots, x_n no sean independientes. Esto nos obliga a desarrollar una nueva teoría probabilística, la teoría de procesos estocásticos, con técnicas diferentes a las utilizadas en la Estadística Matemática elemental. Si suponemos que x_1, \dots, x_n es una realización de variables aleatorias X_1, \dots, X_n , ésta supone que dichas variables son independientes, considerando:

$$\text{Cov}(X_i, X_j) = 0 \quad \forall i, j \quad (4.2)$$

que en nuestro caso no sería debido a la relación que tienen en función del tiempo. Por esta razón, tenemos que desarrollar nuevas técnicas para su estudio.

4.1. Análisis clásico (Determinista)

Si representamos cualquier serie temporal gráficamente, observaremos cosas como una tendencia, periodicidades, datos erróneos...

En el análisis clásico de series temporales, sobre el cual no nos centraremos en este trabajo, consiste en recoger estas observaciones: tendencia, componentes cíclicas, componente estacional y componente residual estacionaria como componentes, las cuales mediante diferentes técnicas de estimación aislaremos para finalmente poder definir:

$$X_i = f(T_i, C_i^{(1)}, \dots, C_i^{(k)}, R_i) \quad (4.3)$$

es decir, escribir nuestros datos como una combinación de dichas componentes.

Normalmente, esta combinación suele ser de forma aditiva

$$X_i = T_i + C_i^{(1)} + \dots + C_i^{(k)} + R_i \quad (4.4)$$

Pero también podemos encontrarlos de forma multiplicativa

$$X_i = T_i \times C_i^{(1)} \times \dots \times C_i^{(k)} \times R_i \quad (4.5)$$

o una combinación de ambos esquemas.

4.2. Procesos estocásticos

Como hemos mencionado anteriormente, nosotros no nos centraremos en el estudio clásico de series temporales, sino que consideraremos nuestros datos como fenómenos que evolucionan en el tiempo. Para ello, los modelos adecuados son los llamados procesos estocásticos.

Definición 4.1 *Un proceso estocástico es una colección de variables aleatorias indexadas por un conjunto \mathbb{T} , que habitualmente representa el tiempo: $\{X_t, t \in \mathbb{T}\}$. El proceso puede ser a tiempo discreto, es decir \mathbb{T} es \mathbb{N} ó \mathbb{Z} , o por otra parte a tiempo continuo, dónde \mathbb{T} es $[0, \infty)$ ó \mathbb{R} .*

Nosotros a partir de ahora veremos las series temporales como una realización de un proceso estocástico.

4.3. Procesos de segundo orden

Definición 4.2 *Un proceso $\{X_t, t \in \mathbb{N} \text{ ó } \mathbb{Z}\}$ es de **segundo orden** si se cumple que*

$$E(X_t^2) < \infty \quad \forall t \quad (4.6)$$

Lo cual garantiza que $\text{var}(X_t)$ está bien definida y es finita, ya que se cumple que:

$$|E(X_t)| \leq E(|X_t|) \leq (E(X_t^2))^{1/2} < \infty \quad (4.7)$$

y, por lo tanto

$$\text{Var}(X_t) := E(X_t^2) - E(X_t)^2 < \infty . \quad (4.8)$$

A partir de ahora, cualquier proceso estocástico será para nosotros de segundo orden.

4.4. Procesos estacionarios

Cada serie temporal tiene su particular forma, cosa que se puede observar al hacer su respectiva gráfica. Necesitamos que los procesos estocásticos sean la representación matemática de algún tipo de regularidad, para que puedan ser utilizados como modelos de nuestras series.

Definición 4.3 Un proceso $\{X_t, t \in \mathbb{Z}\}$ es **estrictamente estacionario** si para cualquier conjunto de variables incluido en este, éstas presentan la misma ley que las que obtenemos desplazándolas una determinada longitud, es decir, que para cualquier t_1, \dots, t_n y l , los vectores

$$(X_{t_1}, \dots, X_{t_n}) \quad (4.9)$$

y

$$(X_{t_1+l}, \dots, X_{t_n+l}) \quad (4.10)$$

presenten la misma ley.

Nota 4.1 Para $n = 1$ implica que todas las X_t tienen la misma ley.

Definición 4.4 Un conjunto es **estacionario** (o estacionario en sentido débil) si todas sus variables aleatorias tienen la misma esperanza y la covarianza que presentan solamente depende de la distancia que las separa, es decir:

- $E[X_t] = \mu \in \mathbb{R}, \quad \forall t \in \mathbb{Z}$
- $Cov(X_t, X_{t+l}) = \gamma(l), \quad \forall t \in \mathbb{Z}$

Nota 4.2 Para $l = 0$ tenemos que $\gamma(0) = Var(X_t)$, $\forall t \geq 1$ y también se cumple que $\gamma(l) = \gamma(-l)$.

A partir de la función $\gamma(l)$, llamada *función de autocovarianza*, obtenemos la función

$$\rho(l) := \frac{\gamma(l)}{\gamma(0)} \quad (4.11)$$

llamada *función de autocorrelación*.

Por lo tanto, para poder analizar nuestras series, necesitaremos hacer un estudio previo y aplicarles las transformaciones que sean necesarias para que presenten las propiedades que acabamos de definir.

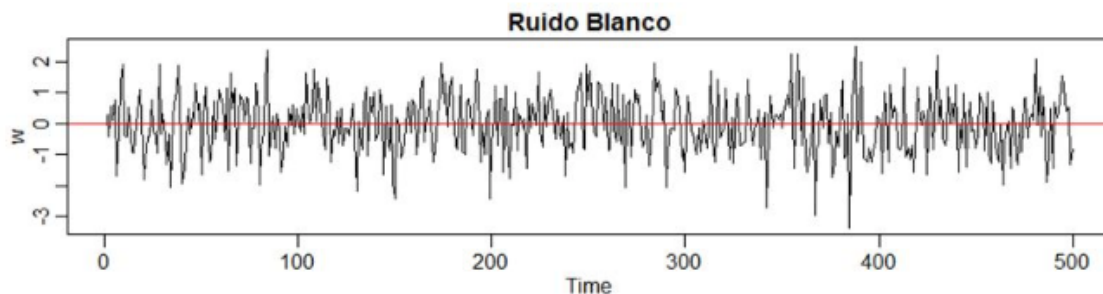
4.5. Ejemplos de procesos estocásticos

4.5.1. Ruido IID

Un proceso $\{X_t, t \geq 1\}$ es un ruido i.i.d si las variables son independientes e idénticamente distribuidas, y con media μ y desviación típica σ . En este caso, la función de autocorrelación toma los valores $\rho(l) = 1$ para $l = 0$ y $\rho(l) = 0 \ \forall l > 0$.

4.5.2. Ruido blanco

Un proceso $\{X_t, t \geq 1\}$ es un ruido blanco si las variables tienen media μ , varianza σ^2 y están incorreladas, es decir, no tienen relación lineal ($Cov(X_t, X_k) = 0 \ \forall t \ \forall k$), pero podrían presentar algún tipo de relación no lineal.



Nota 4.3 *Un ruido i.i.d es un caso particular de ruido blanco.*

4.5.3. Ruido blanco Gaussiano

Un proceso $\{X_t, t \geq 1\}$ es un ruido blanco gaussiano si a parte de ser blanco, todo vector $(X_{t_1}, \dots, X_{t_n})$ tiene ley normal.

Nota 4.4 *Un ruido blanco gaussiano es un caso particular de ruido i.i.d, ya que en este caso la incorrelación implica independencia.*

5. Modelos ARCH y GARCH

5.1. Introducción a los modelos ARCH

Muchas series temporales financieras, como las acciones, tasas de interés, tipos de cambio, etc, tienen unas determinadas características que los procesos lineales clásicos, como los procesos AR o ARMA, no las tienen en cuenta en su modelado.

Uno de estos *stylized facts* que tienen dichas series es la heterocedasticidad condicional, es decir, una variabilidad condicional no contante a lo largo de la serie. Este hecho no se tiene en cuenta en los procesos lineales clásicos, ya que una de las premisas de estos es tener una varianza constante, tanto condicional como incondicional, a lo largo de la serie temporal.

Para solucionar este problema, aparecen modelos en los que la evolución de la volatilidad condicional juega un papel importante. Los modelos más populares son los nombrados *modelos con heterocedasticidad condicional autoregresiva* (ARCH), los cuales tienen en cuenta estas determinadas características.

Los modelos ARCH aparecen en los años 80, propuestos por Robert F. Engle, un economista, estadístico y profesor universitario que mediante estos, dió la posibilidad de poder analizar la volatilidad condicional que presentan la mayoría de las series del mercado financiero, en las cuales aparecen períodos turbulentos, con cambios bruscos, seguidos de períodos de calma con apenas fluctuaciones. Gracias a estos, Engle fue galardonado con el Premio Nobel de Economía el año 2003.

Sin embargo, se han descubierto desventajas relevantes en los modelos ARCH, que han dado lugar a un gran número de extensiones de estos, como los modelos GARCH, EGARCH... Los cuales facilitan el modelado y pueden ajustar mejor los datos.

5.2. Stylized facts

Los *stylized facts* que presentan este tipo de series temporales financieras, y en los modelos nombrados se tienen en cuenta son los siguientes:

Stylized Fact 1 *Determinadas series temporales X_t , como por ejemplo el precio de algunas acciones, no son estacionarias, ya que suelen presentar algún tipo de tendencia.*

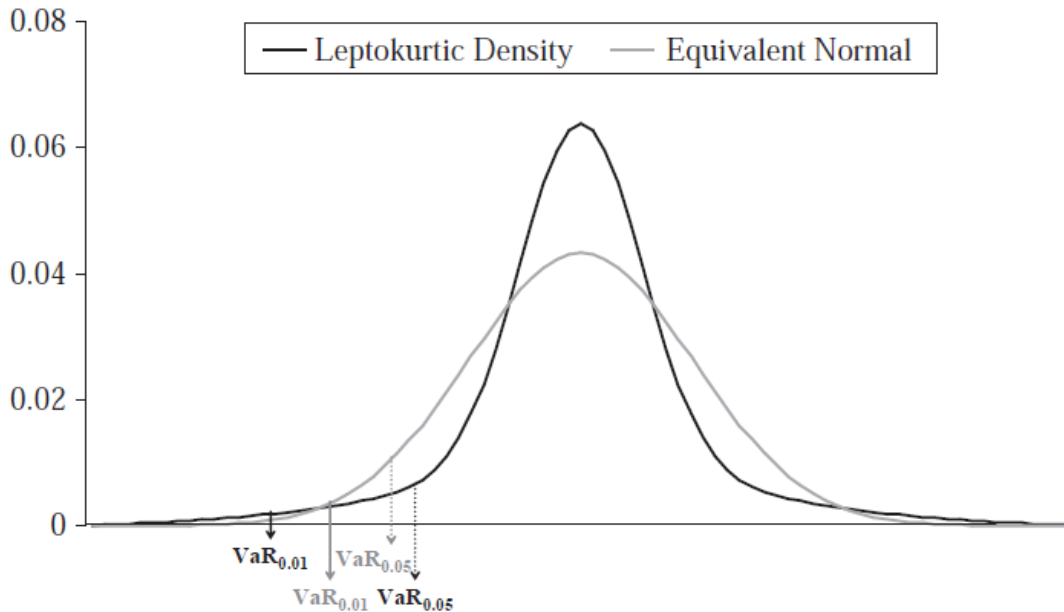
Para solucionar este problema, se tiene que realizar determinadas transformaciones a la serie temporal. Primero se hace el logaritmo de los datos, es decir, $Y_t = \log X_t$, para evitar una volatilidad creciente o decreciente. Después de esto, se toman diferencias a resultados obtenidos, es decir, $R_t = Y_t - Y_{t-1}$. Por lo tanto, la transformación conjunta sería:

$$R_t = \log \frac{X_t}{X_{t-1}} = \log X_t - \log X_{t-1} \quad (5.1)$$

que se llama rendimiento logarítmico.

Stylized Fact 2 *La serie R_t tiene una distribución leptocúrtica, ya que su estimación empírica de la curtosis es principalmente superior a 3, que es la que tiene la distribución normal estándar.*

Un proceso ARCH tiene una curtosis superior a 3, lo que implica una mayor concentración de valores de la variable cerca de la media de la distribución y en las colas, dejando las zonas intermedias a estas menos pobladas, haciendo que su forma tenga colas más gruesas y un centro más apuntado si la comparamos con la distribución normal estándar (mesocúrtica).



Stylized Fact 3 *En los mercados financieros se observa que el proceso R_t es un ruido blanco, ya que la función de autocorrelación de la muestra $\rho(l)$, $l \neq 0$ no es significativamente diferente de 0. A pesar de esto, no es un ruido iid, es decir, no son independientes ya que la función de autocorrelación del cuadrado o del valor absoluto de la muestra del proceso son claramente superiores a 0.*

Por lo tanto, tenemos que $Cov(R_i, R_{i+l}) = 0$, es decir, hay incorrelación. Pero sin embargo, no son independientes ya que, por ejemplo, $Cov(R_i^2, R_{i+l}^2)$ y $Cov(|R_i|, |R_{i+l}|)$ no tienen que ser iguales a 0, ya que los datos no tienen porque ser independientes.

Stylized Fact 4 *La volatilidad de estas series temporales suele formar grupos: Después de grandes (pequeños) cambios, ya sean positivos como negativos, suelen producirse cambios grandes (pequeños). A este efecto se le llama **volatility clustering**.*

A partir de esta propiedad, existe una correlación positiva entre los cuadrados de los rendimientos. Así, como estos están heterocedásticamente condicionados,

$$Var(R_{t+1}|\mathcal{F}_t) \neq Var(R_{t+1}) \quad (5.2)$$

es decir, la variabilidad de los rendimientos depende de sus recientes cambios.

A causa de todas estas características propias que presentan algunas series financieras, se ha tenido la necesidad de utilizar modelos alternativos a los procesos lineales clásicos. Es así como aparecen los procesos ARCH y todas sus respectivas extensiones.

5.3. Modelo ARCH(1): Definición y propiedades

Definición 5.1 Un proceso R_t , $t \in \mathbb{Z}$ es un **ARCH(1)** (Modelo con heterocedasticidad condicional autorregresiva de orden 1) si cumple:

- $E(R_t|\mathcal{F}_{t-1}) = 0$
- $Var(R_t|\mathcal{F}_{t-1}) = \sigma_t^2 = w + \alpha R_{t-1}^2$

con $w > 0$, $\alpha \geq 0$ y $Z_t = R_t/\sigma_t$ donde $E(Z_t) = 0$ y $Var(Z_t) = 1$.

De esta manera, el modelo quedaría representado como:

$$R_t = Z_t \sigma_t = Z_t (w + \alpha R_{t-1}^2)^{1/2} \quad (5.3)$$

A este modelo se le llama **semi-strong ARCH**.

Nota 5.1 Si además se cumple que $Z_t = R_t/\sigma_t$ es i.i.d, se le llama **strong ARCH**.

Frecuentemente asumiremos que Z_t está normalmente distribuido $Z_t \sim N(0, 1)$, lo que significaría que R_t está condicionalmente distribuido como una normal:

$$R_t|\mathcal{F}_{t-1} \sim N(0, \sigma_t^2) \quad (5.4)$$

En este caso los modelos *semi-strong* y *strong* son iguales.

5.3.1. Propiedades del modelo ARCH(1)

Teorema 5.1 Asumiendo que R_t es un proceso *semi-strong ARCH(1)* con $Var(R_t) = \sigma^2 < \infty$, entonces:

$$\sigma^2 = \frac{w}{1 - \alpha} \quad (5.5)$$

Demostración: $\sigma^2 = E(R_t^2) = E(E(R_t^2|\mathcal{F}_{t-1})) = E(\sigma_t^2) = w + \alpha E(R_{t-1}^2) \implies$

Y como se cumple que $E(R_t^2) = E(R_{t-1}^2)$ tenemos que:

$$\implies E(R_t^2) = w + \alpha E(R_t^2) \implies \sigma^2 = w + \alpha \sigma^2 \implies (1 - \alpha)\sigma^2 = w \implies \sigma^2 = \frac{w}{1 - \alpha}$$

Siempre y cuando $\alpha < 1$ (condición necesaria y suficiente para que un proceso *semi-strong* sea estacionario en sentido débil).

□

Teorema 5.2 *Asumiendo que R_t es un proceso semi-strong ARCH(1) con $Var(R_t) = \sigma^2 < \infty$, se deduce que R_t es un ruido blanco.*

Demostración: Por definición, sabemos que $E(R_t|\mathcal{F}_{t-1}) = 0$, lo que nos hace deducir que $E(R_t) = 0$. Por lo tanto, $Cov(R_t, R_{t-l}) = E((R_t - E(R_t))^2(R_{t-l} - E(R_{t-l}))^2) = E(R_t R_{t-l}) = E(E(R_t R_{t-l}|\mathcal{F}_{t-1})) = E(R_{t-l} E(R_t|\mathcal{F}_{t-1})) = 0$. Lo que confirma que R_t es un ruido blanco. \square

Nota 5.2 *Este R_t no es un ruido i.i.d.*

Si se cumple que $Z_t = R_t/\sigma_t$ esta simétricamente distribuido alrededor del zero $E(Z_t) = 0$, entonces todos los momentos impares de R_t son zero y, bajo la suposición de distribución normal estándar (expresión (5.4)), derivamos la existencia de momentos pares superiores, dando lugar a lo siguiente:

Teorema 5.3 *Sea R_t un proceso strong ARCH(1), $Z_t \sim N(0, 1)$ y $E(R_t^4) = c < \infty$. Entonces:*

- $E(R_t^4) = \frac{3w^2}{(1-\alpha)^2} \frac{1-\alpha^2}{1-3\alpha^2}$, con $3\alpha^2 < 1$
- $Kurt(R_t) \geq 3$. Es decir, la distribución incondicional de R_t es **leptocúrtica**.

Demostración: $c = E(R_t^4) = E(E(R_t^4|\mathcal{F}_{t-1})) = E(E(\sigma_t^4 Z_t^4|\mathcal{F}_{t-1})) = E(\sigma_t^4 E(Z_t^4|\mathcal{F}_{t-1})) \implies$

Asumiendo que $\sigma_t^4 = (\sigma_t^2)^2 = (w + \alpha R_t^2)^2$, $E(Z_t^4|\mathcal{F}_{t-1}) = E(Z_t^4) = 3$ (distribución normal)

$$\implies E(R_t^4) = 3E((w + \alpha R_t^2)^2) = 3E(w^2 + 2w\alpha R_t^2 + \alpha^2 R_t^4) = 3(w^2 + 2w\alpha E(R_t^2) + \alpha^2 E(R_t^4)) \implies$$

Como hemos visto en el teorema anterior, $E(R_t^2) = \frac{w}{1-\alpha}$ por lo tanto:

$$\implies E(R_t^4) = 3(w^2 + \frac{2w^2\alpha}{1-\alpha} + \alpha^2 E(R_t^4)) = 3w^2 + \frac{6w^2\alpha}{1-\alpha} + 3\alpha^2 E(R_t^4) \implies$$

$$\implies c = 3w^2 + \frac{6w^2\alpha}{1-\alpha} + 3\alpha^2 c \implies (1 - 3\alpha^2)c = 3w^2 + \frac{6w^2\alpha}{1-\alpha} \implies$$

$$\implies (1 - 3\alpha^2)c = \frac{3w^2(1-\alpha) + 6w^2\alpha}{1-\alpha} = \frac{3w^2 - 3w^2\alpha + 6w^2\alpha}{1-\alpha} = \frac{3w^2 + 3w^2\alpha}{1-\alpha} \implies$$

$$\begin{aligned} \Rightarrow c &= \frac{3w^2(1+\alpha)}{(1-3\alpha^2)(1-\alpha)} \cdot \frac{1-\alpha}{1-\alpha} = \frac{3w^2(1-\alpha^2)}{(1-3\alpha^2)(1-\alpha)^2} = \frac{3w^2}{(1-\alpha)^2} \cdot \frac{1-\alpha^2}{1-3\alpha^2} \Rightarrow \\ \Rightarrow c &= \frac{3w^2}{(1-\alpha)^2} \cdot \frac{1-\alpha^2}{1-3\alpha^2}, \text{ con } 3\alpha^2 < 1 \end{aligned}$$

Por lo respecta a la curtosis de R_t :

$$Kurt(R_t) = \frac{E(R_t^4)}{E(R_t^2)^2} = \frac{\frac{3w^2}{(1-\alpha)^2} \cdot \frac{1-\alpha^2}{1-3\alpha^2}}{\frac{w^2}{(1-\alpha)^2}} = 3 \cdot \frac{1-\alpha^2}{1-3\alpha^2} \geq 3, \text{ ya que } \frac{1-\alpha^2}{1-3\alpha^2} \geq 1 \text{ si } 3\alpha^2 < 1$$

□

Por lo tanto, $\forall \alpha \neq 0$ la distribución incondicional es *leptocúrtica*, es decir, la curva de su distribución es más alta en el medio y las colas son más anchas que las de la distribución normal, rasgo que nos encontramos normalmente en los mercados financieros.

Para sacar conclusiones empíricamente sobre el espesor de las colas de la distribución incondicional, podemos asumir que esta es del tipo Pareto para valores altos de x , es decir:

$$P(x) = P(X_t > x) \sim kx^{-a}, \quad a > 0. \quad (5.6)$$

Para estimar el parámetro a hay múltiples opciones, como utilizar la log función de $P(x)$:

$$\log P(x) \approx \log k - a \log x \quad (5.7)$$

poniendo los estadísticos en orden $X_{(1)} > X_{(2)} > \dots > X_{(n)}$ y estimando la probabilidad $P(x)$ para cada $x = X_{(i)}$ usando la frecuencia relativa:

$$\frac{\#\{t; X_t \geq X_{(i)}\}}{n} = \frac{i}{n}. \quad (5.8)$$

Así, sustituyendo este estimador en la expresión 5.7, podemos estimar el parámetro a a partir de la regresión:

$$\log \frac{i}{n} \approx \log k - a \log x \quad (5.9)$$

utilizando el método de mínimos cuadrados, solamente con una pequeña parte de los datos (los de orden superior) para estimar la regresión, ya que esta distribución solo es apropiada en las colas.

Hemos visto que podemos representar al proceso R_t utilizando la expresión:

$$R_t = Z_t \sigma_t = Z_t (w + \alpha R_{t-1}^2)^{1/2}. \quad (5.10)$$

Sin embargo, en el caso del proceso *strong* ARCH(1), podemos representar R_t^2 como una sucesión de retardos del cuadrado ruido blanco Z_t , el cual en este tipo de procesos es independiente e idénticamente distribuido y $E(Z_t) = 0$ y $Var(Z_t) = E(Z_t^2) = 1$.

Teorema 5.4 Sea R_t un proceso *strong* ARCH(1) con $Var(R_t) = \sigma^2 < \infty$. Entonces,

$$R_t^2 = w \sum_{k=0}^{\infty} \alpha^k \prod_{j=0}^k Z_{t-j}^2 \quad (5.11)$$

y esta suma converge en L_1 , es decir $E(|R_t^2 - w \sum_{k=0}^m \alpha^k \prod_{j=0}^k Z_{t-j}^2|) \rightarrow 0$ cuando $m \rightarrow \infty$.

Demostración: Utilizando que $R_t^2 = Z_t^2 \sigma_t^2$ y que $\sigma_t^2 = w + \alpha R_{t-1}^2$ tenemos:

$$R_t^2 = Z_t^2 \sigma_t^2 = Z_t^2 (w + \alpha R_{t-1}^2) = w Z_t^2 + \alpha Z_t^2 R_{t-1}^2$$

volviendo a utilizar las dos expresiones iniciales:

$$\begin{aligned} R_t^2 &= w Z_t^2 + \alpha R_{t-1}^2 Z_t^2 = w Z_t^2 + \alpha Z_t^2 Z_{t-1}^2 \sigma_{t-1}^2 = w Z_t^2 + \alpha Z_t^2 Z_{t-1}^2 (w + \alpha R_{t-2}^2) = \\ &= w Z_t^2 + w \alpha Z_t^2 Z_{t-1}^2 + \alpha^2 Z_t^2 Z_{t-1}^2 R_{t-2}^2 = w Z_t^2 + w \alpha Z_t^2 Z_{t-1}^2 + w \alpha^2 Z_t^2 Z_{t-1}^2 Z_{t-2}^2 + \alpha^3 Z_t^2 Z_{t-1}^2 Z_{t-2}^2 R_{t-3}^2 = \\ &= w (Z_t^2 + \alpha Z_t^2 Z_{t-1}^2 + \alpha^2 Z_t^2 Z_{t-1}^2 Z_{t-2}^2) + \alpha^3 Z_t^2 Z_{t-1}^2 Z_{t-2}^2 R_{t-3}^2 = \dots = w \sum_{k=0}^{\infty} \alpha^k \prod_{j=0}^k Z_{t-j}^2 \implies \\ &\implies R_t^2 = w \sum_{k=0}^{\infty} \alpha^k \prod_{j=0}^k Z_{t-j}^2 \end{aligned}$$

Y por lo respecta a la convergencia en L_1 , en la iteración m tenemos:

$$\begin{aligned} R_t^2 &= w (Z_t^2 + \alpha Z_t^2 Z_{t-1}^2 + \dots + \alpha^m Z_t^2 Z_{t-1}^2 \dots Z_{t-m}^2) + \alpha^{m+1} Z_t^2 Z_{t-1}^2 \dots Z_{t-m}^2 R_{t-m-1}^2 = \\ &= w \sum_{k=0}^m \alpha^k \prod_{j=0}^k Z_{t-j}^2 + \alpha^{m+1} R_{t-m-1}^2 \prod_{j=0}^m Z_{t-j}^2 \end{aligned}$$

Por lo tanto si lo aplicamos a la convergencia en L_1 :

$$\begin{aligned} E(|R_t^2 - w \sum_{k=0}^m \alpha^k \prod_{j=0}^k Z_{t-j}^2|) &= E(\alpha^{m+1} R_{t-m-1}^2 \prod_{j=0}^m Z_{t-j}^2) = \alpha^{m+1} E(R_{t-m-1}^2 \prod_{j=0}^m Z_{t-j}^2) = \\ &= \alpha^{m+1} E(R_{t-m-1}^2) \rightarrow 0 \quad \text{si } m \rightarrow \infty \\ &\quad (Z_t \text{ i.i.d y } E(Z_t^2) = 1 \quad \forall t) \end{aligned}$$

□

Teorema 5.5 Sea R_t un proceso strong ARCH(1) estacionario, con $E(R_t^4) = c < \infty$ y $Z_t \sim N(0, 1)$. Entonces se cumple que,

- $R_t^2 = w \sum_{k=0}^{\infty} \alpha^k \prod_{j=0}^k Z_{t-j}^2$ y esta suma converge en L_2 .
- $\eta_t = \sigma_t^2(Z_t^2 - 1)$ es un ruido blanco.
- R_t^2 es un proceso AR(1) con $R_t^2 = w + \alpha R_{t-1}^2 + \eta_t$

Demostración: La primera expresión se demuestra de la misma manera que en el Teorema 5.4. Por lo que a la convergencia en L_2 respecta:

$$\begin{aligned}
 E((R_t^2 - w \sum_{k=0}^m \alpha^k \prod_{j=0}^k Z_{t-j}^2)^2) &= E((\alpha^{m+1} R_{t-m-1}^2 \prod_{j=0}^m Z_{t-j}^2)^2) = \\
 &= \alpha^{2(m+1)} E(R_{t-m-1}^4) E(\prod_{j=0}^m Z_{t-j}^4) = \\
 &= \alpha^{2(m+1)} 3^{m+1} c \longrightarrow 0 \text{ si } m \longrightarrow \infty \\
 &(E(R_{t-m-1}^4) = c \text{ y } Z_t \sim N(0, 1), \text{ independiente con } Kurt(Z_t) = E(Z_t^4) = 3)
 \end{aligned}$$

Ahora, comprobamos que $\eta_t = \sigma_t^2(Z_t^2 - 1)$ es un ruido blanco:

$$\begin{aligned}
 E(\eta_t) &= E(\sigma_t^2(Z_t^2 - 1)) = E(\sigma_t^2)E(Z_t^2 - 1) = E(\sigma_t^2)(E(Z_t^2) - 1) = 0 \quad (E(Z_t^2) = Var(Z_t^2) = 1) \\
 Var(\eta_t) &= E(\eta_t^2) = E(\sigma_t^4(Z_t^2 - 1)^2) = E(\sigma_t^4)E((Z_t^2 - 1)^2) = E(\sigma_t^4)E(Z_t^4 - 2Z_t^2 + 1) = \\
 &= 2E((w + \alpha R_{t-1}^2)^2) = 2(w^2 + 2\alpha w E(R_{t-1}^2) + \alpha^2 E(R_{t-1}^4)) = constante \\
 &(E(Z_t^4 - 2Z_t^2 + 1) = 3 - 2 + 1 = 2; \quad E(R_{t-1}^2) = \sigma^2 \text{ y } E(R_{t-1}^4) = c)
 \end{aligned}$$

Es decir, la varianza es independiente de t .

$$Cov(\eta_t, \eta_{t+l}) = E(\sigma_t^2(Z_t^2 - 1)\sigma_{t+l}^2(Z_{t+l}^2 - 1)) = E(\sigma_t^2(Z_t^2 - 1)\sigma_{t+l}^2)E((Z_{t+l}^2 - 1)) = 0 \quad \forall s \neq 0$$

Por lo tanto, η_t es un ruido blanco. Por último, la expresión que aparece al final del teorema saldría de:

$$R_t^2 = \sigma_t^2 Z_t^2 = \sigma_t^2(1 + Z_t^2 - 1) = \sigma_t^2 + \sigma_t^2(Z_t^2 - 1) = w + \alpha R_{t-1}^2 + \eta_t$$

Vemos entonces que R_t^2 depende de una contante w , un ruido blanco η_t y de su término anterior R_{t-1}^2 , convirtiéndolo en un proceso AR(1).

□

5.3.2. Estimación de los modelos ARCH(1)

En el Teorema 5.5 hemos visto que podemos representar un proceso ARCH(1) como un AR(1) en X_t^2 . Utilizando esta propiedad, podemos usar el estimador de Yule-Walker para estimar los parámetros w y α del proceso:

$$\hat{\alpha}^{(0)} = \frac{\sum_{t=2}^n (R_t^2 - \hat{w}^{(0)})(R_{t-1}^2 - \hat{w}^{(0)})}{\sum_{t=2}^n (R_t^2 - \hat{w}^{(0)})^2} \quad (5.12)$$

con $\hat{w}^{(0)} = \frac{\sum_{t=1}^n R_t^2}{n}$. Pero como la distribución de R_t^2 no es normal, este estimador es ineficiente.

El método de estimación habitual para los parámetros de los modelos ARCH se basa en el método de la máxima verosimilitud, asumiendo que R_t tiene condicionalmente una distribución normal. Por lo tanto tenemos que:

$$p(R_t | \mathcal{F}_{t-1}) = \frac{1}{\sqrt{2\pi}\sigma_t} \exp\left(-\frac{R_t^2}{2\sigma_t^2}\right) \quad (5.13)$$

Usando la la función de log-verosimilitud $l(w, \alpha)$ podemos escribir la función con los parámetros a estimar:

$$\begin{aligned} l(w, \alpha) &= \sum_{t=2}^n l_t(w, \alpha) + \log p_R(R_1) \\ &= \sum_{t=2}^n \log p(R_t | \mathcal{F}_{t-1}) + \log p_R(R_1) \\ &= \sum_{t=2}^n \log \left(\frac{1}{\sqrt{2\pi}\sigma_t} \exp\left(-\frac{R_t^2}{2\sigma_t^2}\right) \right) + \log p_R(R_1) \\ &= \sum_{t=2}^n \log \left(\frac{1}{\sqrt{2\pi}\sigma_t} \right) - \sum_{t=2}^n \frac{R_t^2}{2\sigma_t^2} + \log p_R(R_1) \end{aligned} \quad (5.14)$$

Tenemos que $\sum_{t=2}^n \log \left(\frac{1}{\sqrt{2\pi}\sigma_t} \right) = \sum_{t=2}^n (\log 1 - \log (2\pi)^{1/2} - \log \sigma_t) =$

$$= -\frac{n-1}{2} \log(2\pi) - \frac{1}{2} \sum_{t=2}^n \log (w + \alpha R_{t-1}^2)$$

Y también que $\sum_{t=2}^n \frac{R_t^2}{2\sigma_t^2} = \sum_{t=2}^n \frac{R_t^2}{2(w + \alpha R_{t-1}^2)} = \frac{1}{2} \sum_{t=2}^n \frac{R_t^2}{w + \alpha R_{t-1}^2}$

Por lo tanto la expresión anterior quedaría como:

$$l(w, \alpha) = -\frac{n-1}{2} \log(2\pi) - \frac{1}{2} \sum_{t=2}^n \log(w + \alpha R_{t-1}^2) - \frac{1}{2} \sum_{t=2}^n \frac{R_t^2}{w + \alpha R_{t-1}^2} + \log p_R(R_1) \quad (5.15)$$

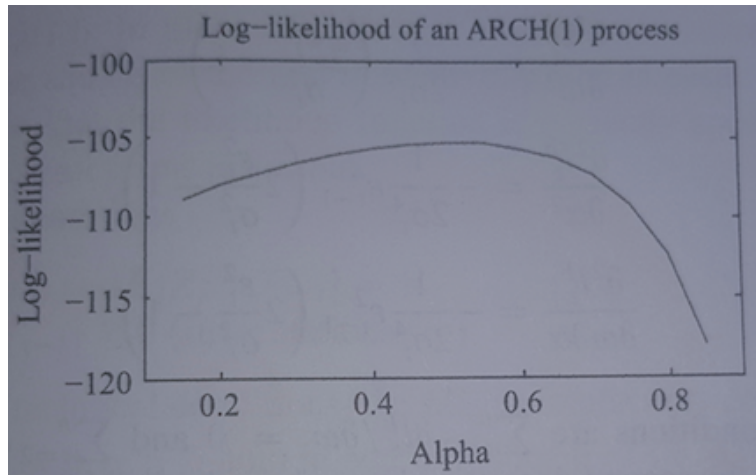
Donde p_R es la densidad marginal estacionaria de R_t , pero debido a que esta expresión es desconocida en los modelos ARCH, no podemos calcularla. Por lo que usaremos la función de verosimilitud $l^b = \log p(R_n, \dots, R_2 | R_1)$, donde $\log p_R(R_1)$ desaparece, y para n grandes las diferencias con l no son relevantes:

$$l^b(w, \alpha) = \sum_{t=2}^n l_t(w, \alpha) = -\frac{n-1}{2} \log(2\pi) - \frac{1}{2} \sum_{t=2}^n \log(w + \alpha R_{t-1}^2) - \frac{1}{2} \sum_{t=2}^n \frac{R_t^2}{w + \alpha R_{t-1}^2} \quad (5.16)$$

A partir de esta función, podemos encontrar el α que haga l^b máximo, y a partir de este, y de la expresión de la varianza no condicional del proceso, podemos encontrar w , ya que se cumple:

$$\sigma^2 = \frac{w}{1-\alpha} \implies w = (1-\alpha)\sigma^2$$

Para encontrar el valor óptimo de la función log-verosímil, podemos analizar su gráfica, buscando el α que nos daría el máximo de esta:



Este es un ejemplo de una función log-verosímil de un proceso ARCH(1), tomada con 100 muestras, en la cual se observa que el valor α óptimo es 0,5 (extraída del libro *Statistics of Financial Markets*).

La precisión del estimador esta determinada por la segunda derivada de la función de verosimilitud en el punto optimo. Para ello, se debe calcular las primeras y segundas derivadas parciales de la función de verosimilitud:

$$\frac{\partial l^b}{\partial w} = \frac{1}{2\sigma_t^2} \left(\frac{R_t^2}{\sigma_t^2} - 1 \right) \quad \frac{\partial l^b}{\partial \alpha} = \frac{1}{2\sigma_t^2} R_{t-1}^2 \left(\frac{R_t^2}{\sigma_t^2} - 1 \right) \quad (5.17)$$

$$\frac{\partial^2 l^b}{\partial w^2} = -\frac{1}{2\sigma_t^4} \left(2\frac{R_t^2}{\sigma_t^2} - 1 \right) \quad \frac{\partial^2 l^b}{\partial \alpha^2} = -\frac{1}{2\sigma_t^4} R_{t-1}^4 \left(2\frac{R_t^2}{\sigma_t^2} - 1 \right) \quad \frac{\partial^2 l^b}{\partial w \partial \alpha} = -\frac{1}{2\sigma_t^4} R_{t-1}^2 \left(2\frac{R_t^2}{\sigma_t^2} - 1 \right) \quad (5.18)$$

Sabemos que las condiciones de primer orden son que $\sum_{t=2}^n \frac{\partial l^b}{\partial w} = 0$ y $\sum_{t=2}^n \frac{\partial l^b}{\partial \alpha}$, y la esperanza del valor de la segunda derivada se tiene que calcular de la misma forma.

Asumiendo que $E(Z_t^2) = E((R_t/\sigma_t)^2) = 1$, se cumple que $(2\frac{R_t^2}{\sigma_t^2} - 1) = 1$. Por lo tanto:

$$E\left(\frac{\partial^2 l^b}{\partial w^2}\right) = -\frac{1}{2}E\left(\frac{1}{\sigma_t^4}\right) \quad (5.19)$$

Y la esperanza de $\frac{1}{\sigma_t^4}$ se estima consistentemente con $\frac{1}{n-1} \sum_{t=2}^n \frac{1}{(w + \alpha R_{t-1}^2)^2}$.

De esta manera el estimador de la esperanza del valor de la segunda derivada serían:

$$\hat{E}\left(\frac{\partial^2 l^b}{\partial w^2}\right) = -\frac{1}{2(n-1)} \sum_{t=2}^n \frac{1}{\sigma_t^4} \quad \hat{E}\left(\frac{\partial^2 l^b}{\partial \alpha^2}\right) = -\frac{1}{2(n-1)} \sum_{t=2}^n \frac{R_{t-1}^4}{\sigma_t^4} \quad (5.20)$$

5.4. Modelo ARCH(q): Definición y propiedades

Cuando la varianza condicional del modelo ARCH depende de más de un retardo, es decir, también se tienen en cuenta los anteriores a este para estimar dicha varianza, se extiende el modelo tomando el nombre de $ARCH(q)$, donde q es el número de retardos que se tienen en cuenta.

Definición 5.2 *El proceso R_t , $t \in \mathbb{Z}$ es un **ARCH(q)** (Modelo con heterocedasticidad condicional autoregresiva de orden q) si cumple:*

- $E(R_t | \mathcal{F}_{t-1}) = 0$
- $Var(R_t | \mathcal{F}_{t-1}) = \sigma_t^2 = w + \alpha_1 R_{t-1}^2 + \dots + \alpha_q R_{t-q}^2 = w + \sum_{j=1}^q \alpha_j R_{t-j}^2$

con $w > 0$, $\alpha_1 \geq 0, \dots, \alpha_q \geq 0$ y $Z_t = R_t / \sigma_t$ donde $E(Z_t) = 0$ y $Var(Z_t) = 1$.

De esta manera, el modelo quedaría representado como:

$$R_t = Z_t \sigma_t = Z_t (w + \alpha_1 R_{t-1}^2 + \dots + \alpha_q R_{t-q}^2)^{1/2} \quad (5.21)$$

A este modelo se le llama **semi-strong ARCH**.

Nota 5.3 *Si además se cumple que $Z_t = R_t / \sigma_t$ es i.i.d, se le llama **strong ARCH**.*

Ahora, la varianza condicional σ_t^2 es una función lineal del cuadrado de los q retardos.

5.4.1. Propiedades del modelo ARCH(q)

Teorema 5.6 Sea R_t un proceso semi-strong ARCH(q) con $\text{Var}(R_t) = \sigma^2 < \infty$, entonces:

$$\sigma^2 = \frac{w}{1 - \alpha_1 - \dots - \alpha_q} \quad (5.22)$$

Con $\alpha_1 + \dots + \alpha_q < 1$.

Demostración: $\sigma^2 = E(R_t^2) = E(E(R_t^2 | \mathcal{F}_{t-1})) = E(\sigma_t^2) = w + \alpha_1 E(R_{t-1}^2) + \dots + \alpha_q E(R_{t-q}^2) \implies$

Y como se cumple que $E(R_t^2) = E(R_{t-1}^2) = \dots = E(R_{t-q}^2)$ tenemos que:

$$\implies E(R_t^2) = w + \alpha_1 E(R_t^2) + \dots + \alpha_q E(R_t^2) \implies \sigma^2 = w + \alpha_1 \sigma^2 + \dots + \alpha_q \sigma^2 \implies$$

$$\implies (1 - \alpha_1 - \dots - \alpha_q) \sigma^2 = w \implies \sigma^2 = \frac{w}{1 - \alpha_1 - \dots - \alpha_q}$$

□

Si en cambio $\alpha_1 + \dots + \alpha_q \geq 1$, entonces la varianza incondicional no existiría y el proceso no sería estacionario en covarianza.

Igual que en los modelos ARCH(1), en su extension con q retardos, también hay una expresión en la que R_t^2 se comporta como un proceso AR:

Teorema 5.7 Sea R_t un proceso semi-strong ARCH(q) con $E(R_t^4) = c < \infty$. Entonces se cumple que:

- $\eta_t = \sigma_t^2(Z_t^2 - 1)$ es un ruido blanco.
- R_t^2 es un proceso AR(q) con $R_t^2 = w + \sum_{i=1}^q \alpha_i R_{t-i}^2 + \eta_t$

Demostración: $E(\eta_t) = E(\sigma_t^2(Z_t^2 - 1)) = E(\sigma_t^2)E(Z_t^2 - 1) = E(\sigma_t^2)(E(Z_t^2) - 1) = 0$

$$\text{Var}(\eta_t) = E(\eta_t^2) = E(\sigma_t^4(Z_t^2 - 1)^2) = E(\sigma_t^4)E((Z_t^2 - 1)^2) = E(\sigma_t^4)E(Z_t^4 - 2Z_t^2 + 1) =$$

$$= 2E((w + \sum_{j=1}^q \alpha_j R_{t-j}^2)^2) = \text{constante}$$

ya que en esta expresión aparecen esperanzas cuadráticas y cuárticas de R_t , todas ellas constantes.

Es decir, la varianza es independiente de t .

$$Cov(\eta_t, \eta_{t+l}) = E(\sigma_t^2(Z_t^2 - 1)\sigma_{t+l}^2(Z_{t+l}^2 - 1)) = E(\sigma_t^2(Z_t^2 - 1)\sigma_{t+l}^2)E((Z_{t+l}^2 - 1)) = 0 \quad \forall s \neq 0$$

Por lo tanto, η_t es un ruido blanco. Por último, la expresión que aparece al final del teorema saldría de:

$$R_t^2 = \sigma_t^2 Z_t^2 = \sigma_t^2(1 + Z_t^2 - 1) = \sigma_t^2 + \sigma_t^2(Z_t^2 - 1) = w + \sum_{i=1}^q \alpha_i R_{t-i}^2 + \eta_t$$

Vemos entonces que R_t^2 depende de una constante w , un ruido blanco η_t y de sus q retardos anteriores $R_{t-1}^2, \dots, R_{t-q}^2$, convirtiéndolo en un proceso AR(q).

□

Uno de los problemas que tienen los modelos ARCH(q) es el uso de grandes q , es decir, de una gran cantidad de retardos, lo que hace perder influencia a la volatilidad reciente. Además, que los parámetros del modelo sean estimados bajo determinadas condiciones para la estacionariedad y que estos tengan que ser estrictamente positivos, hacen que sea difícil su obtención.

Una posibilidad para reducir el numero de parámetros es añadirles unos pesos lineales decrecientes a los retardos, es decir:

$$\sigma_t^2 = w + \alpha \sum_{i=1}^q \kappa_i R_{t-i}^2 \tag{5.23}$$

$$\kappa_i = \frac{2(q+1-i)}{q(q+1)}$$

De esta manera, solamente se estiman dos parámetros: w y α

5.4.2. Estimación de los modelos ARCH(q)

Igual que en los modelos ARCH(1), la estimación de los ARCH(q) se hace por el método de máxima verosimilitud, en el cual su función sería:

$$l^b(\theta) = \sum_{t=q+1}^n l^t(\theta) = -\frac{n-1}{2} \log(2\pi) - \frac{1}{2} \sum_{t=2}^n \log(\sigma_t^2) - \frac{1}{2} \sum_{t=q+1}^n \frac{R_t^q + 1}{\sigma_t^2} \quad (5.24)$$

Donde $\theta = (w, \alpha_1, \dots, \alpha_q)^T$.

Aunque podemos encontrar el valor optimo de los parámetros observando la gráfica como en el caso del ARCH(1), es complicado para un gran número de parámetros, ya que cada uno de estos implica aumentar una dimensión del gráfico. Sin embargo, es posible maximizar dicha función numéricamente, utilizando el algoritmo *score*, en el cual es necesario calcular la primera y segunda derivada de la función de verosimilitud condicional respecto a los parámetros de esta. La primera derivada quedará como:

$$\frac{\partial l^b}{\partial \theta} = \frac{1}{2\sigma_t^2} \frac{\partial \sigma_t^2}{\partial \theta} \left(\frac{R_t^q}{\sigma_t^2} - 1 \right) \quad (5.25)$$

Con $\frac{\partial \sigma_t^2}{\partial \theta} = (1, R_{t-1}^2, \dots, R_{t-q}^2)^T$

Donde la condición es que $\sum_{t=q+1}^n \frac{\partial l^b}{\partial \theta} = 0$. Las condiciones de la segunda derivada, así como las propiedades asintóticas del estimador quedan explicadas más adelante, en el modelo ARCH generalizado.

5.5. ARCH generalizado: GARCH

En muchas series, como por ejemplo en las financieras, el número de retardos a utilizar es muy elevado, lo cual dificulta su estimación, ya que ello llevaría a un considerable número de iteraciones para alcanzar la solución del sistema planteado, o incluso llegar a no encontrarla debido a las condiciones que presentan sus parámetros. Engle ya propuso ciertas restricciones, como la de añadir pesos a los retardos con tal de reducir el número de parámetros a estimar, aunque estas no recogen cualquier caso.

Es por eso que Tim Peter Bollerslev, un economista danés que actualmente ejerce de profesor en la Universidad Duke en Estados Unidos, propone *la generalización del modelo con heterocedasticidad condicional autorregresiva (GARCH)*, en el cual la varianza condicional no solo depende de los cuadrados de las perturbaciones, como en el modelo ARCH de Engle, sino que además también depende de las varianzas condicionales de períodos anteriores, es decir, de σ_t^2 pasados.

5.6. GARCH(p,q): Definición y propiedades

Definición 5.3 *Un proceso R_t , $t \in \mathbb{Z}$ es un **GARCH(p,q)** (Modelo con heterocedasticidad condicional autorregresiva generalizado) si cumple:*

- $E(R_t | \mathcal{F}_{t-1}) = 0$
- $Var(R_t | \mathcal{F}_{t-1}) = \sigma_t^2 = w + \alpha_1 R_{t-1}^2 + \dots + \alpha_q R_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + \dots + \beta_p \sigma_{t-p}^2 =$
 $= w + \sum_{i=1}^q \alpha_i R_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2$

Donde las condiciones suficientes, aunque no necesarias, para que se cumpla que $\sigma_t^2 > 0$, es decir, $P(\sigma_t^2 > 0) = 1$ son que $w > 0$, $\alpha_1 \geq 0, \dots, \alpha_q \geq 0$, $\beta_1 \geq 0, \dots, \beta_p \geq 0$ y $Z_t = R_t / \sigma_t$ con $E(Z_t) = 0$ y $Var(Z_t) = 1$. Por ejemplo, para el modelo GARCH(1,2):

$$\begin{aligned}
 \sigma_t^2 &= w + \alpha_1 R_{t-1}^2 + \alpha_2 R_{t-2}^2 + \beta_1 \sigma_{t-1}^2 = \\
 &= w + \alpha_1 R_{t-1}^2 + \alpha_2 R_{t-2}^2 + \beta_1 (w + \alpha_1 R_{t-2}^2 + \alpha_2 R_{t-3}^2 + \beta_1 \sigma_{t-2}^2) = \\
 &= w + \alpha_1 R_{t-1}^2 + \alpha_2 R_{t-2}^2 + \beta_1 w + \alpha_1 \beta_1 R_{t-2}^2 + \alpha_2 \beta_1 R_{t-3}^2 + \beta_1^2 \sigma_{t-2}^2 = \\
 &= \dots = w(1 + \beta_1 + \beta_1^2 + \dots) + \alpha_1 \sum_{i=0}^{\infty} \beta_1^i R_{t-i-1}^2 + \alpha_2 \sum_{j=0}^{\infty} \beta_1^j R_{t-j-2}^2 = \\
 &= \frac{w}{1 - \beta_1} + \alpha_1 R_{t-1}^2 + (\alpha_1 \beta_1 + \alpha_2) \sum_{j=0}^{\infty} \beta_1^j R_{t-j-2}^2
 \end{aligned}$$

Por lo tanto, para el modelo $GARCH(1,2)$, $0 \leq \beta_1 < 1$, $w > 0$, $\alpha_1 \geq 0$ y $\alpha_1\beta_1 + \alpha_2 \geq 0$ son condiciones necesarias y suficientes siempre y cuando $\sum_{j=0}^{\infty} \beta_1^j R_{t-j-2}^2$ sea convergente.

Como $Z_t = R_t/\sigma_t$ el modelo se representaría de la forma:

$$R_t = Z_t \sigma_t = Z_t (w + \alpha_1 R_{t-1}^2 + \dots + \alpha_q R_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + \dots + \beta_p \sigma_{t-p}^2)^{1/2} \quad (5.26)$$

A este modelo se le llama **semi-strong GARCH**.

Nota 5.4 Si además se cumple que $Z_t = R_t/\sigma_t$ es i.i.d, se le llama **strong GARCH**.

Ahora, la varianza condicional σ_t^2 es una función lineal del cuadrado de los q retardos y de las p varianzas condicionales anteriores.

5.6.1. Propiedades del modelo GARCH(p,q)

Teorema 5.8 Sea R_t un proceso semi-strong $GARCH(p,q)$ con $Var(R_t) = \sigma^2 < \infty$, entonces:

$$\sigma^2 = \frac{w}{1 - \sum_{i=1}^q \alpha_i - \sum_{j=1}^p \beta_j} \quad (5.27)$$

Con $\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j < 1$.

Demostración: $\sigma^2 = E(R_t^2) = E(E(R_t^2|\mathcal{F}_{t-1})) = E(\sigma_t^2) =$

$$= w + \sum_{i=1}^q \alpha_i E(R_{t-i}^2) + \sum_{j=1}^p \beta_j E(\sigma_{t-j}^2) \implies$$

Y como se cumple que $E(R_t^2) = E(\sigma_t^2)$ y que $E(R_t^2) = E(R_{t-1}^2) = \dots = E(R_{t-q}^2)$ tenemos que:

$$\implies E(R_t^2) = w + \sum_{i=1}^q \alpha_i E(R_t^2) + \sum_{j=1}^p \beta_j E(\sigma_t^2) \implies \sigma^2 = w + \sum_{i=1}^q \alpha_i \sigma^2 + \sum_{j=1}^p \beta_j \sigma^2 \implies$$

$$\implies (1 - \sum_{i=1}^q \alpha_i - \sum_{j=1}^p \beta_j) \sigma^2 = w \implies \sigma^2 = \frac{w}{1 - \sum_{i=1}^q \alpha_i - \sum_{j=1}^p \beta_j}$$

□

Teorema 5.9 Sea R_t un proceso semi-strong GARCH(1,1) con $\text{Var}(R_t) = \sigma^2 < \infty$ y $Z_t \sim N(0,1)$. Entonces se cumple que el momento de cuarto orden de R_t es constante ($E(R_t^4) < \infty$) siempre y cuando $3\alpha_1 + 2\alpha_1\beta_1 + \beta_1^2 < 1$. En ese caso, su kurtosis se define como:

$$\text{Kurt}(R_t) = \frac{E(R_t^4)}{E(R_t^2)^2} = 3 + \frac{6\alpha_1^2}{1 - \beta_1^2 - 2\alpha_1\beta_1 - 3\alpha_1^2} \quad (5.28)$$

Demostración: Tenemos que $E(R_t^4) = E(Z_t^4 \sigma_t^4) = 3E((w + \alpha_1 R_{t-1}^2 + \beta_1 \sigma_{t-1}^2)^2) =$

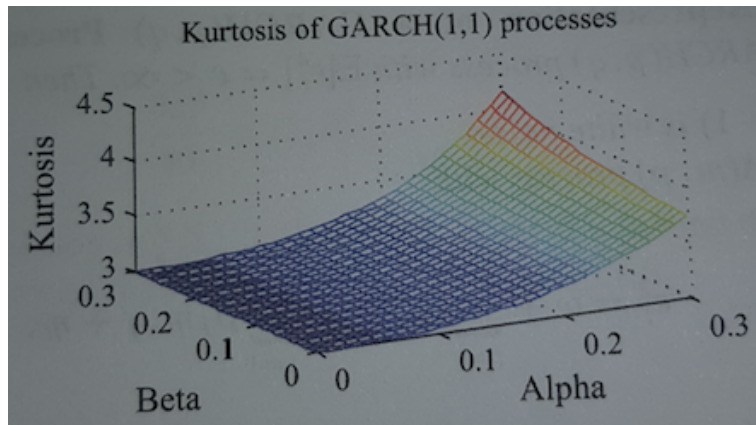
$$= 3 \frac{w^2}{(1 - \beta_1^2 - 2\alpha_1\beta_1 - 3\alpha_1^2)} \frac{(1 - \beta_1^2 - 2\alpha_1\beta_1 - 3\alpha_1^2) + 2\alpha_1^2}{(1 - \alpha_1 - \beta_1)^2}$$

Aparte de comprobar la estacionariedad de R_t . Entonces la kurtosis de R_t sería:

$$\begin{aligned} \text{Kurt}(R_t) &= \frac{E(R_t^4)}{E(R_t^2)^2} = 3 \frac{w^2}{(1 - \beta_1^2 - 2\alpha_1\beta_1 - 3\alpha_1^2)} \frac{(1 - \beta_1^2 - 2\alpha_1\beta_1 - 3\alpha_1^2) + 2\alpha_1^2}{(1 - \alpha_1 - \beta_1)^2} \frac{(1 - \alpha_1 - \beta_1)^2}{w^2} = \\ &= 3 \frac{(1 - \beta_1^2 - 2\alpha_1\beta_1 - 3\alpha_1^2) + 2\alpha_1^2}{(1 - \beta_1^2 - 2\alpha_1\beta_1 - 3\alpha_1^2)} = \\ &= 3 \frac{(1 - \beta_1^2 - 2\alpha_1\beta_1 - 3\alpha_1^2)}{(1 - \beta_1^2 - 2\alpha_1\beta_1 - 3\alpha_1^2)} + \frac{6\alpha_1^2}{(1 - \beta_1^2 - 2\alpha_1\beta_1 - 3\alpha_1^2)} = 3 + \frac{6\alpha_1^2}{1 - \beta_1^2 - 2\alpha_1\beta_1 - 3\alpha_1^2} \end{aligned}$$

□

Vemos que la distribución de R_t es leptocúrtica, ya que $\forall \alpha_1 > 0$, $\text{Kurt}(R_t) > 3$. Solo cuando $\alpha_1 = 0$ la kurtosis es 3, es decir, cuando desaparece la heterocedasticidad condicional y tiene lugar un ruido blanco Gaussiano. Si observamos el gráfico que nos da la kurtosis en función de α y β :



Vemos que si fijamos un α , al aumentar β aumenta la kurtosis mucho más lentamente que en el caso contrario, es decir, si fijamos β y aumentamos α (extraída del libro *Statistics of Financial Markets*).

Teorema 5.10 Sea R_t un proceso semi-strong GARCH(p, q) con $E(R_t^4) = c < \infty$. Entonces se cumple que:

- $\eta_t = \sigma_t^2(Z_t^2 - 1)$ es un ruido blanco.
 - R_t^2 es un proceso ARMA(q, p) con $R_t^2 = w + \sum_{i=1}^m \gamma_i R_{t-i}^2 + \sum_{j=1}^p \beta_j \eta_{t-j} + \eta_t$
- Donde $m = \max(p, q)$, $\gamma_i = \alpha_i + \beta_i \alpha_i = 0$ cuando $i > q$, y $\beta_i = 0$ cuando $i > p$.

Demostración: La demostración es la misma que nos encontramos en el Teorema 5.5

□

Por lo que si R_t sigue un proceso GARCH, R_t^2 sigue un proceso ARMA, en el cual sabemos que si todas las raíces del polinomio $(1 - \beta_1 - \dots - \beta_p)$ están fuera del círculo unidad, el proceso es invertible y se puede escribir como un AR(∞). De esta manera, podemos sacar conclusiones de el orden de (p, q) a partir de los modelos ARMA.

5.6.2. Estimación de los modelos GARCH(p, q)

Si nos basamos en la representación ARMA de los procesos GARCH, podemos considerar otra vez los estimadores de Yule-Walker, igual que se hacia en la estimación del modelo ARCH(1). Sin embargo, de la misma manera que en el proceso anterior, estos no son eficientes, ya que la distribución de R_t^2 no es normal. Por lo que para estimar los parámetros del modelo, se utilizara de nuevo el método de la máxima verosimilitud.

La función de verosimilitud tiene la misma expresión que la 5.24 que nos encontramos en la estimación de los modelos ARCH(q), con la diferencia de que ahora $\theta = (w, \alpha_1, \dots, \alpha_q, \beta_1, \dots, \beta_p)^T$.

El parámetro w lo escogemos utilizando la expresión de la varianza incondicional, ya que esta es constante y $\sigma^2 = \frac{w}{1 - \sum_{i=1}^q \alpha_i - \sum_{j=1}^p \beta_j} \implies w = \sigma^2(1 - \sum_{i=1}^q \alpha_i - \sum_{j=1}^p \beta_j)$.

Las primeras derivadas parciales serían:

$$\frac{\partial l^b}{\partial \theta} = \frac{1}{2\sigma_t^2} \frac{\partial \sigma_t^2}{\partial \theta} \left(\frac{R_t^q}{\sigma_t^2} - 1 \right) \quad (5.29)$$

$$\text{Con } \frac{\partial \sigma_t^2}{\partial \theta} = \vartheta_t + \sum_{j=1}^p \frac{\partial \sigma_{t-j}^2}{\partial \theta} \quad \text{y} \quad \vartheta_t = (1, R_{t-1}^2, \dots, R_{t-q}^2, \sigma_{t-1}^2, \dots, \sigma_{t-p}^2)^T,$$

y las condiciones de primer orden son que $\sum_{t=q+1}^n \frac{\partial l_t}{\partial \theta} = 0$.

La matriz correspondiente a las segundas derivadas tiene la forma:

$$\frac{\partial^2 l_t(\theta)}{\partial \theta \partial \theta^T} = \frac{1}{2\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \theta} \frac{\partial \sigma_t^2}{\partial \theta^T} - \frac{1}{2\sigma_t^2} \frac{\partial^2 \sigma_t^2(\theta)}{\partial \theta \partial \theta^T} - \frac{R_t^2}{\sigma_t^6} \frac{\partial \sigma_t^2}{\partial \theta} \frac{\partial \sigma_t^2}{\partial \theta^T} + \frac{R_t^2}{2\sigma_t^4} \frac{\partial^2 \sigma_t^2(\theta)}{\partial \theta \partial \theta^T} \quad (5.30)$$

Asumiendo que $E(Z_t|\mathcal{F}_{t-1}) = 0$, $E(Z_t^2|\mathcal{F}_{t-1}) = 1$, que R_t es estrictamente estacionario y otras condiciones técnicas que aseguran la consistencia del estimador de máxima verosimilitud. Si además $E(Z_t^4|\mathcal{F}_{t-1}) < \infty$, se cumple que $\hat{\theta}$ sigue una distribución asintóticamente normal:

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{L} N_{p+q-1}(0, J^{-1} I J^{-1}) \quad (5.31)$$

Con $I = E\left(\frac{\partial l_t(\theta)}{\partial \theta} \frac{\partial l_t(\theta)}{\partial \theta^T}\right)$ y $J = -E\left(\frac{\partial^2 l_t(\theta)}{\partial \theta \partial \theta^T}\right)$.

Teorema 5.11 Si $Z_t \sim N(0, 1)$, entonces $I = J$.

Demostración: Si calculo dichas esperanzas:

$$J = -E\left(\frac{\partial^2 l_t(\theta)}{\partial \theta \partial \theta^T}\right) = E\left(\frac{1}{2\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \theta} \frac{\partial \sigma_t^2}{\partial \theta^T}\right)$$

$$\begin{aligned} I &= E\left(\frac{\partial l_t(\theta)}{\partial \theta} \frac{\partial l_t(\theta)}{\partial \theta^T}\right) = E\left(\frac{1}{2\sigma_t^2} \frac{\partial \sigma_t^2}{\partial \theta} \left(\frac{R_t^q}{\sigma_t^2} - 1\right) \frac{1}{2\sigma_t^2} \frac{\partial \sigma_t^2}{\partial \theta^T} \left(\frac{R_t^q}{\sigma_t^2} - 1\right)\right) = E\left(\frac{1}{4\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \theta} \frac{\partial \sigma_t^2}{\partial \theta^T} \left(\frac{R_t^q}{\sigma_t^2} - 1\right)^2\right) = \\ &= E\left(\frac{1}{4\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \theta} \frac{\partial \sigma_t^2}{\partial \theta^T} (Z_t^2 - 1)^2\right) = E\left(\frac{1}{4\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \theta} \frac{\partial \sigma_t^2}{\partial \theta^T} (Z_t^4 - 2Z_t^2 + 1)\right) = \\ &= E\left(\frac{1}{4\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \theta} \frac{\partial \sigma_t^2}{\partial \theta^T}\right) E(Z_t^4 - 2Z_t^2 + 1) = E\left(\frac{1}{4\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \theta} \frac{\partial \sigma_t^2}{\partial \theta^T}\right) (E(Z_t^4) - 2E(Z_t^2) + 1) = \\ &= E\left(\frac{1}{4\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \theta} \frac{\partial \sigma_t^2}{\partial \theta^T}\right) (3 - 2 + 1) = 2E\left(\frac{1}{4\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \theta} \frac{\partial \sigma_t^2}{\partial \theta^T}\right) = E\left(\frac{1}{2\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \theta} \frac{\partial \sigma_t^2}{\partial \theta^T}\right) \end{aligned}$$

Asumiendo que $Z_t \sim N(0, 1)$ y por lo tanto $E(Z_t^2|\mathcal{F}_{t-1}) = 1$ y $E(Z_t^4|\mathcal{F}_{t-1}) = Kurt(Z_t|\mathcal{F}_{t-1}) = 3$

□

Por lo que, si Z_t está especificada correctamente, $I = J$ y la varianza asintótica pasaría de ser $J^{-1} I J^{-1}$ a ser J^{-1} , es decir, la inversa de la matriz J , llamada *Matriz de Información de Fisher*.

5.7. Extensiones de los modelos GARCH

En los modelos GARCH estándar se asume que los errores positivos y negativos influyen simétricamente sobre la volatilidad, es decir, tanto la información buena como la mala causan el mismo efecto sobre la volatilidad del modelo. Este supuesto en la práctica suele ser erróneo, como por ejemplo en el rendimiento de las acciones, ya que en estos la volatilidad aumenta más después de malas noticias que de buenas. A este efecto se le llama **Leverage Effect** y aparece por primera vez a través de Fischer Sheffey Black, un economista americano.

Esta propiedad es comprensible, ya que rendimientos negativos implican una alta proporción de deuda, que queda reflejado con un valor de mercado reducido de la empresa, lo que provoca una mayor volatilidad. Sin embargo, aunque los rendimientos positivos causan aumentos en menor escala, estos hacen también hacen crecer la volatilidad.

Por lo que desde el punto de vista empírico, la volatilidad no reacciona simétricamente al signo de los choques y es por eso que se han sugerido extensiones parametrizadas del modelo GARCH estándar. Dos de las más importantes son el modelo **GARCH Exponencial (EGARCH)** y el **Threshold GARCH (TGARCH)** (Modelo GARCH con umbral).

5.7.1. Modelo GARCH Exponencial

Considerando Z_t una serie i.i.d de variables aleatorias estandarizadas con esperanza 0 y varianza 1, el modelo **GARCH Exponencial general (EGARCH)**, dado por Nelson, tiene como varianza condicional la siguiente expresión:

$$\log \sigma_t^2 = w_t + \sum_{k=1}^{\infty} \beta_k g(Z_{t-k}) \quad (5.32)$$

Donde w_t y β_k son coeficientes deterministas y

$$g(Z_t) = \theta Z_t + \gamma(|Z_t| - E(|Z_t|)). \quad (5.33)$$

De donde vemos que $E(g(Z_t)) = 0$.

El modelo EGARCH muestra varias diferencias respecto al modelo GARCH estándar:

- La volatilidad del modelo EGARCH es una función multiplicativa explícita de las innovaciones retardadas, a diferencia del modelo GARCH, donde es una función aditiva de los errores retardados R_t , lo que causa una complicada dependencia funcional.
- La volatilidad puede reaccionar asimétricamente ante términos de error positivos y negativos.
- Para las distribuciones generales de Z_t las restricciones de parámetros para strong y covarianza-estacionariedad coinciden.
- Los parámetros que aparecen en el modelo ya no tienen que ser positivos, dando más libertad a sus estimaciones.

La función $g(\cdot)$ es lineal por partes y contiene dos parámetros, los cuales controlan los efectos de tamaño y signo de los choques de la volatilidad. El término $\gamma(|Z_t| - E(|Z_t|))$ controla el efecto asimétrico, como puede ser el *Leverage Effect*, y determina el efecto de tamaño y el término θZ_t es un efecto ARCH típico y determina el efecto del signo. Normalmente el parámetro γ es positivo y θ es negativo.

Para estimar los modelos EGARCH, se utiliza una representación ARMA(p,q), es decir:

$$\Delta(L) \log \sigma_t^2 = w + \Psi(L)g(Z_t) \quad (5.34)$$

Con $\Delta(L)$ y $\Psi(L)$ como los polinomios retardados de orden p y q, respectivamente.

La ventaja de los modelos EGARCH es que los parámetros no tienen ninguna restricción, pero tienen por contra que no hay unas propiedades claras para los estimadores de cuasi máxima verosimilitud.

5.7.2. Modelo Threshold ARCH

La idea de los modelos **Threshold ARCH (TARCH)** es dividir la distribución de las innovaciones de la desviación y la varianza condicionales en intervalos disjuntos, para luego aproximarlas a una función lineal por partes. Si solo hay dos intervalos, entonces la división se hace en el 0, separando las innovaciones positivas de las negativas para la influencia en la volatilidad. En este caso, el modelo TARCH de orden q se puede escribir como:

$$\sigma_t^\delta = \omega + \sum_{i=1}^q \alpha_i R_{t-i}^\delta + \sum_{j=1}^q \alpha_j^- R_{t-j}^\delta \mathbf{1}(R_{t-j} < 0) \quad (5.35)$$

Donde $\mathbf{1}(\cdot)$ es la función indicador, y para $\delta = 1$ tenemos la expresión de la desviación estándar condicional, y para $\delta = 2$ la de la varianza condicional.

Extendiendo este modelo añadiéndole las desviaciones estándar condicionales retardadas (varianzas respectivamente) como regresores, aparece el llamado modelo **TGARCH**.

En lugar de una función lineal por partes, se puede utilizar una función gradual, constante por partes, como modelo de la volatilidad. Así aparece el modelo **threshold ARCH cualitativo (QTARCH)** el cual se escribe, para el orden 1:

$$y_t = \sum_{j=1}^J m_j \mathbf{1}(y_{t-1} \in A_j) + \sum_{j=1}^J m_j \mathbf{1}(y_{t-1} \in A_j) Z_t \quad (5.36)$$

Donde A_j , $j = 1, \dots, J$ es una partición de la distribución de las innovaciones, y m_j y s_j son escalares. La esperanza y desviación estándar condicionales se modelan mediante una función por partes.

Los modelos threshold tienen una transición entre dos regímenes con fuertes desniveles cuando la innovación generadora cruza un valor umbral. Para tener una transición más suave, la función del indicador puede sustituirse por una función continua que tienda a cero en los valores que están cerca de R_{t-i} , y a uno para valores más lejanos. Normalmente se escoge la función logística, quedando el modelo de la siguiente manera:

$$\sigma_t^\delta = \omega + \sum_{i=1}^q \alpha_i R_{t-i}^\delta + \sum_{j=1}^q \alpha_j^- R_{t-j}^\delta F_\gamma(R_{t-j}) \quad (5.37)$$

Donde $F_\gamma(u) = \frac{1}{1 + \exp(-\gamma u)}$, con el parámetro $\gamma > 0$. Cuando $\gamma \rightarrow \infty$, nos da la función indicador.

Finalmente hay que mencionar un modelo que puede replicar las asimetrías, llamado **ARCH de Asimetría de Potencia (APARCH)**:

$$\sigma_t^\delta = \omega + \sum_{i=1}^q \alpha_i (|R_{t-i}| - \gamma_i R_{t-i})^\delta + \sum_{j=1}^p \beta_j \sigma_{t-j}^\delta \quad (5.38)$$

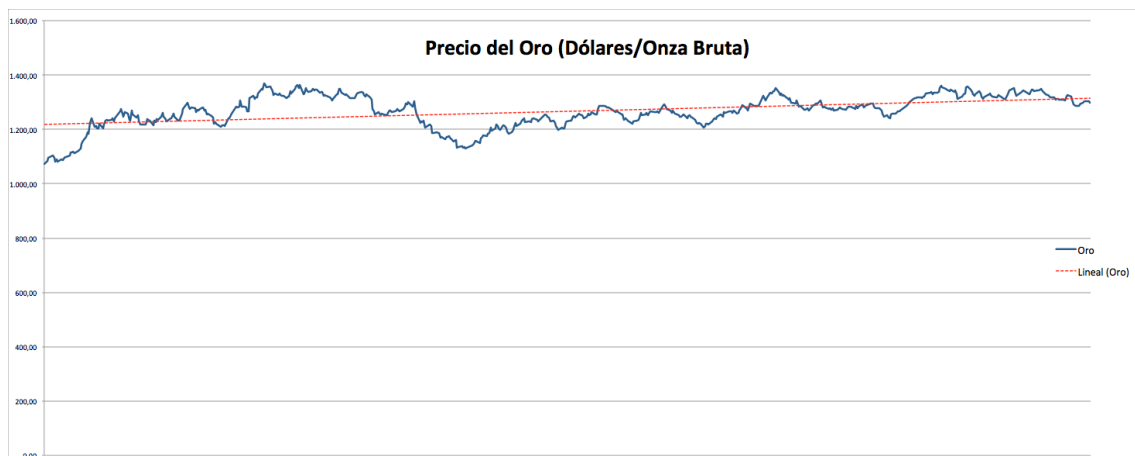
6. Aplicaciones de los modelos ARCH y GARCH

En este apartado se intentará llevar a la práctica lo comentado anteriormente mediante el programa **R**, comprobando si diferentes series temporales financieras muestran los rasgos característicos que se tienen en cuenta en los modelos con heterocedasticidad condicional autorregresiva y, posteriormente, se verá cual de estos se ajusta mejor cada una de estas series.

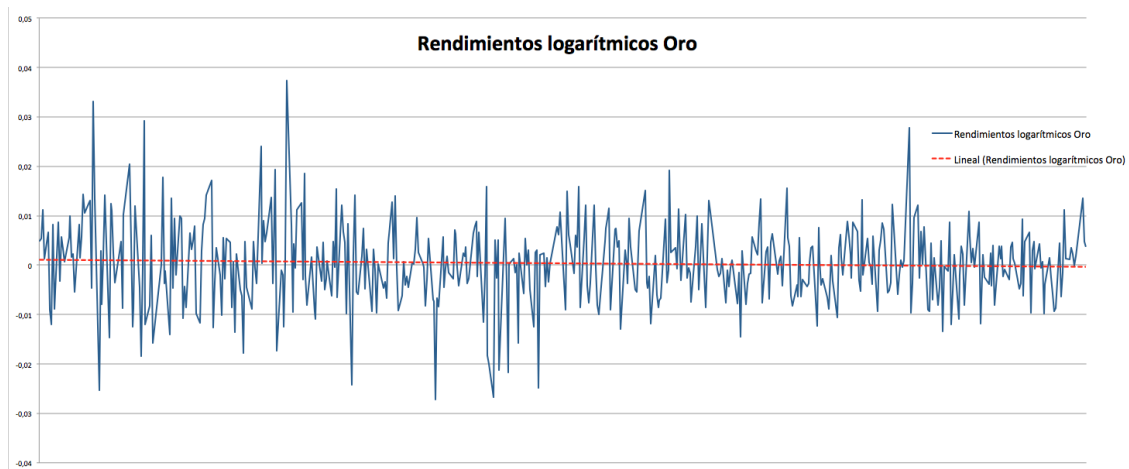
Estas han sido obtenidas de la web <http://www.infobolsa.es/> y todas son series financieras diarias, que van desde enero del 2016 hasta mayo de 2018.

6.1. Precio del oro (Dólar/Onza)

Esta serie temporal define el precio diario del oro, en dolares/onza bruta, de los años 2016 y 2017. Su representación gráfica es la siguiente:

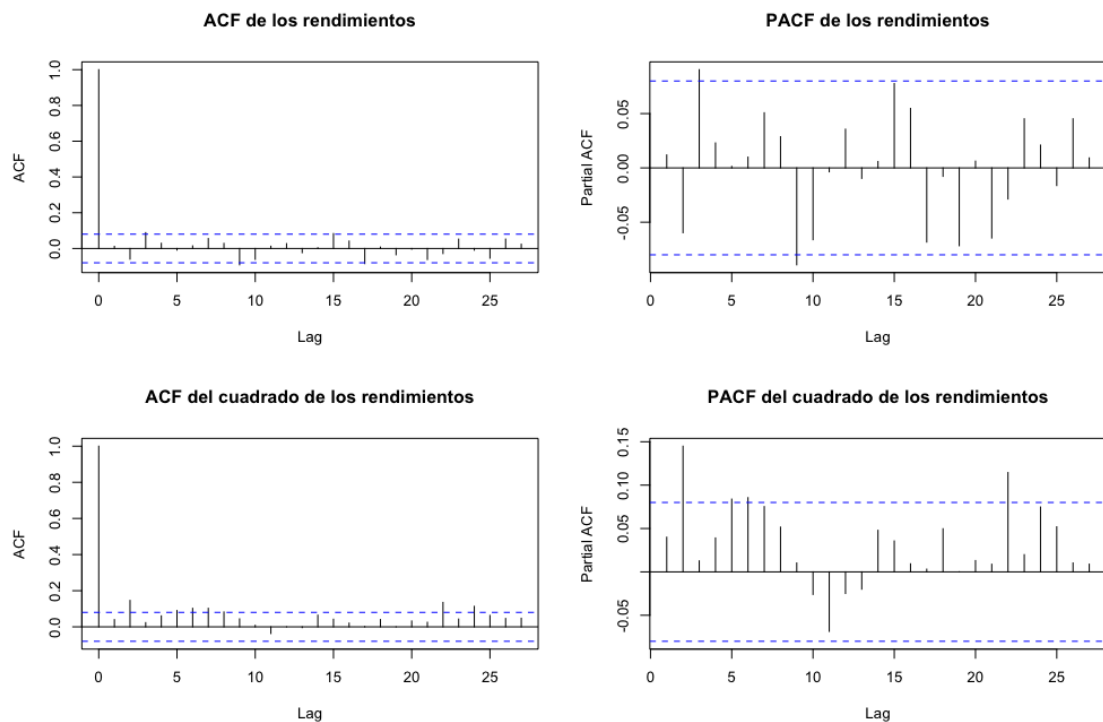


Para poder analizarla con los modelos con heterocedasticidad condicional autorregresiva, tenemos que transformarla a rendimientos logarítmicos. La serie trasformada tendría la forma:

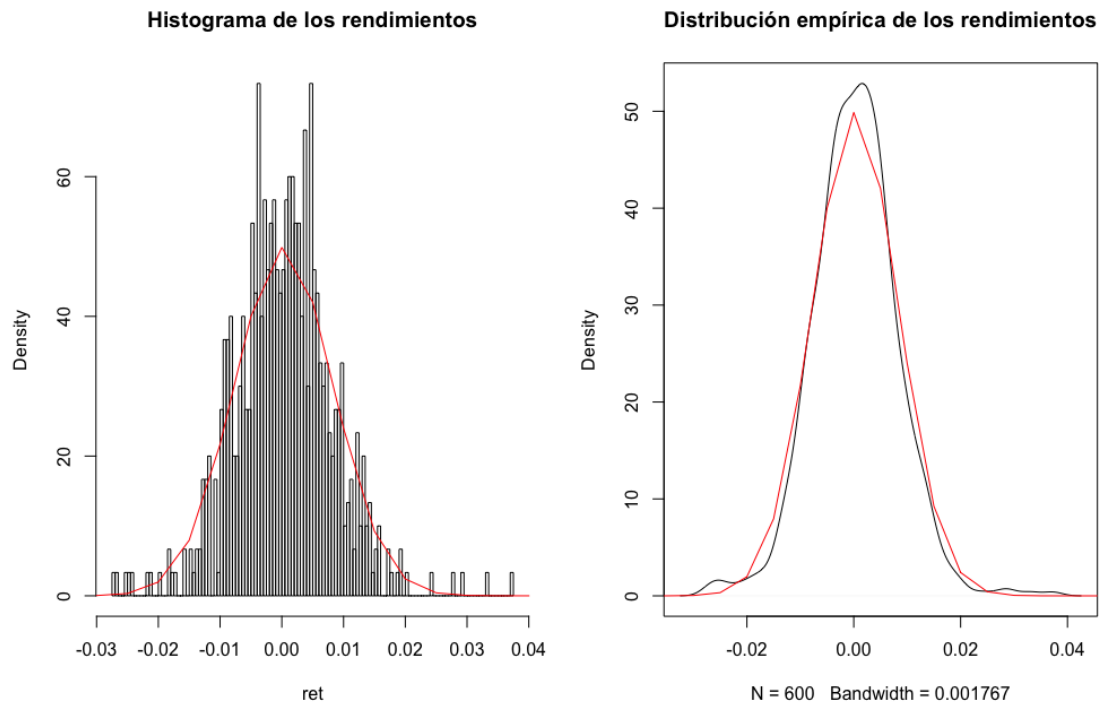


Vemos que esta serie no tiene tendencia, ya que los valores oscilan alrededor del 0, y que la variabilidad va cambiando, por lo que podría modelarse con alguno de los modelos estudiados.

Ahora, antes de mirar qué modelo de los comentados se ajusta más a esta serie, comprobaremos si cumple alguno de los *Stylized Facts* propios de estos modelos. Primero de todo, miraremos las gráficas de las funciones de autocorrelación y autocorrelación parcial, tanto de la serie como la del cuadrado de esta, para comprobar que no existe correlación de primer orden, pero si de orden cuadrático, uno de los rasgos de estos modelos:



Se aprecia que los momento de primer orden no están correlacionados. En cambio, aunque no se aprecia mucho, parece que los de segundo orden si que estaán más correlacionados si los comparamos con los primeros. Ahora, miraremos tanto el histograma como la distribución empírica de la serie, y la compararemos con la distribución normal, con la media y la varianza incondicional de nuestros datos:

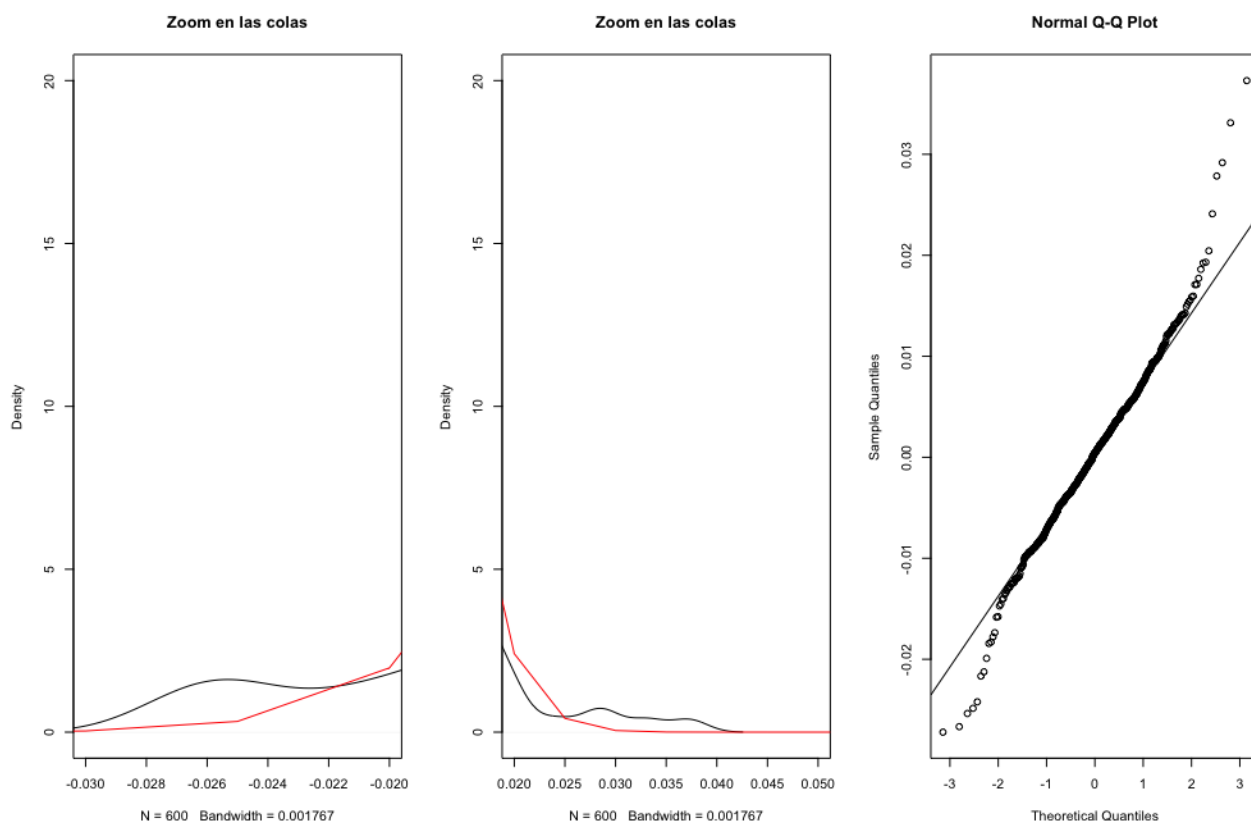


Se puede apreciar, aunque no de una manera muy notable, que si comparamos la distribución empírica de los rendimientos con la distribución normal, esta tiene una curtosis ligeramente superior, y unas colas un poco mas gruesas. Si hacemos la comprobación numérica de la kurtosis, utilizando el paquete *moments* de **R**:

```
> kurtosis(ret)
V1
4.737541
```

Observamos que esta es de 4.737541, un poco superior a la de la distribución normal estándar, que es de 3.

Si finalmente hacemos zoom en las colas y el QQ-plot de la distribución normal:



Se confirma que las colas son un poco más gruesas, aunque se ajusta bastante bien a la distribución normal. Por lo que en se concluiría que tiene rasgos de los modelos ARCH, aunque no son muy notables.

Si hacemos el test de Ljung-Box, en el cual si rechazamos la hipótesis nula, nos indica que los datos no se distribuyen de forma independiente, y por lo tanto nos da un rasgo de los modelos ARCH, el resultado es el siguiente:

Box-Pierce test

```
data: ts^2
X-squared = 0.98135, df = 1, p-value = 0.3219
```

No es significativo al 5 %, por lo tanto aceptaríamos la hipótesis nula que indica que los datos se distribuyen de forma independiente. De todos modos, buscaremos el modelo que mejor ajuste a esta serie.

Mediante el paquete *fGarch* de **R** y utilizando el criterio de AIC, nos quedaremos con el modelo que mejor se ajuste. De todos los modelos vistos, el que mejor se ajusta es un GARCH(1,1):

Error Analysis:

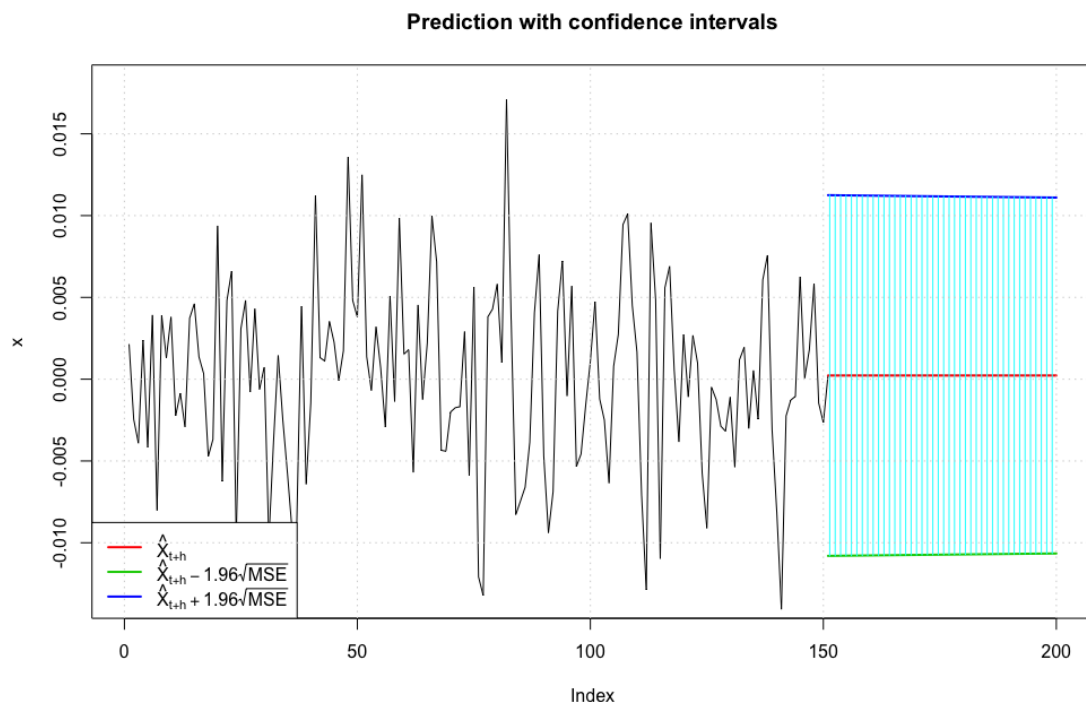
	Estimate	Std. Error	t value	Pr(> t)
mu	2.239e-04	3.033e-04	0.738	0.46047
omega	5.204e-08	1.615e-07	0.322	0.74723
alpha1	1.518e-02	5.185e-03	2.928	0.00342 **
beta1	9.826e-01	5.745e-03	171.046	< 2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Vemos que solamente los dos últimos parámetros salen significativos al 5 %, quedando el modelo de la forma:

$$\sigma_t^2 = 0,01518R_{t-1}^2 + 0,9826\sigma_{t-1}^2, \quad R_t = \sigma_t Z_t \quad \text{con} \quad Z_t \sim N(0,1)$$

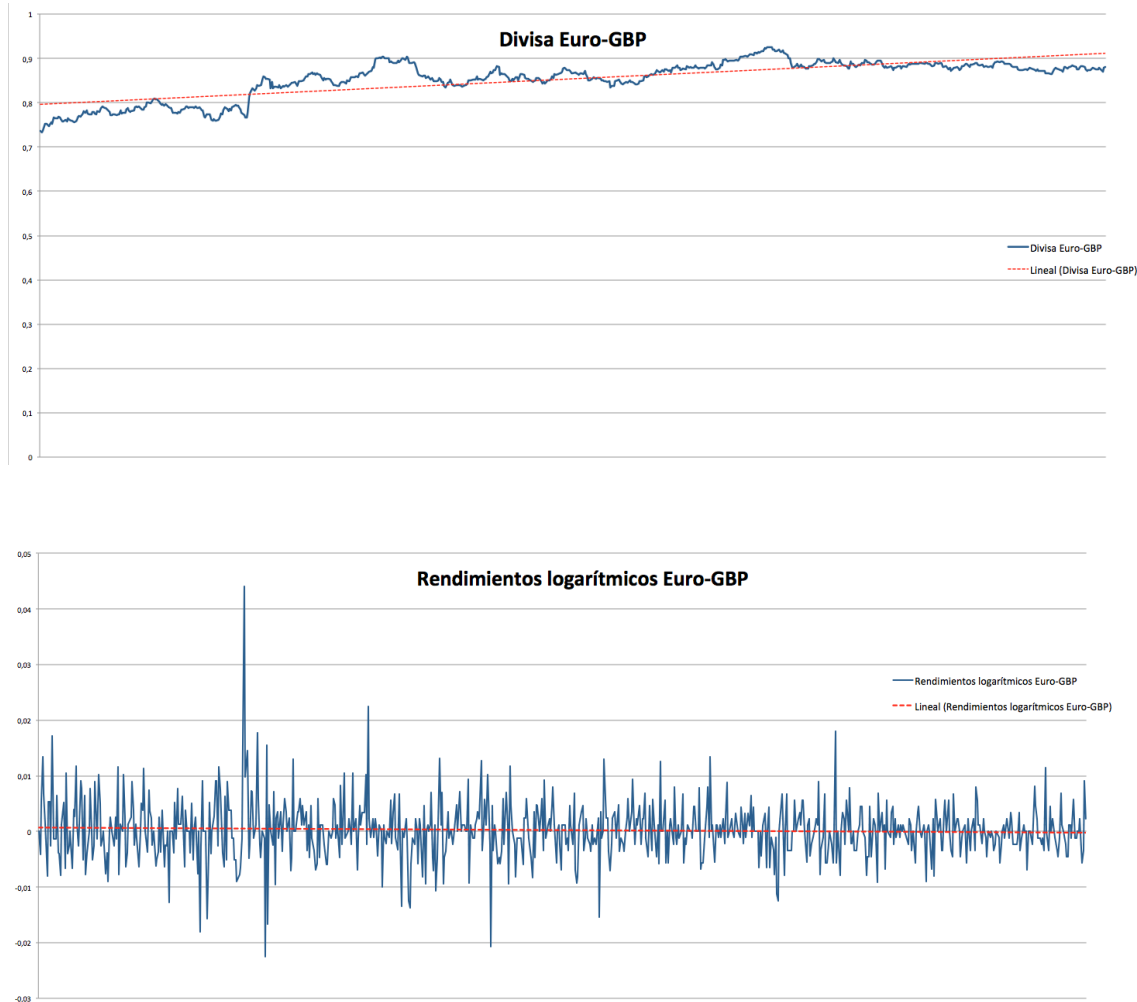
Y la gráfica, añadiéndole las 50 predicciones futuras, quedaría como:



Donde la raya roja indica la media, y la azul y la verde la varianza máxima y mínima alcanzable, respectivamente.

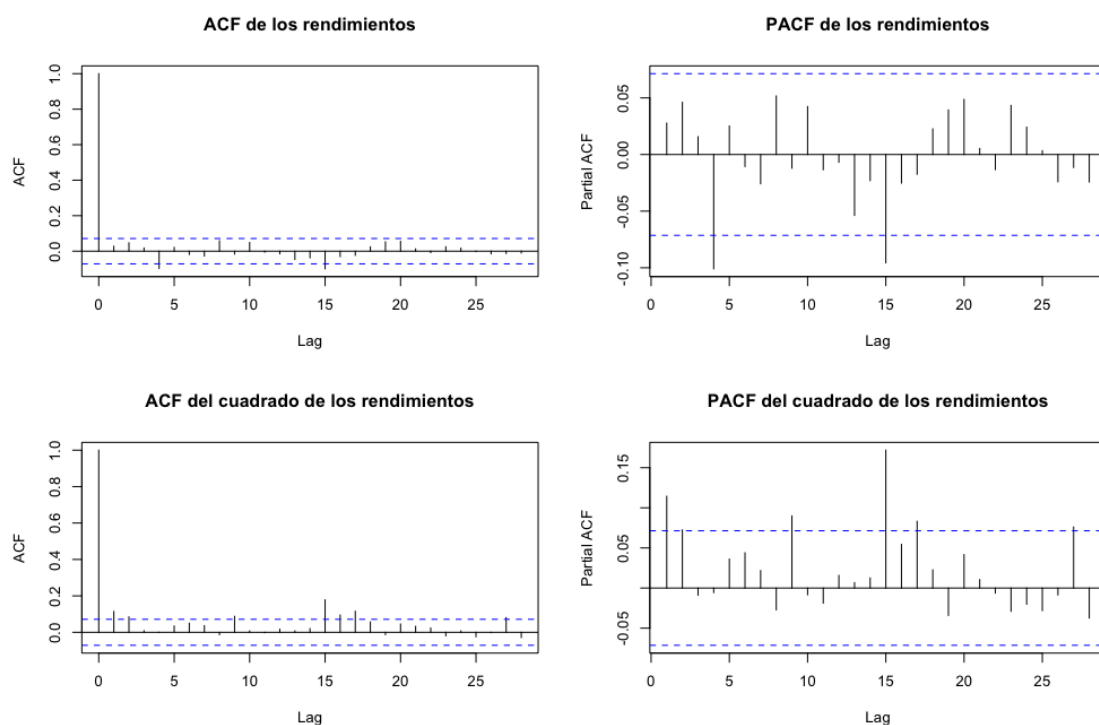
6.2. Divisa Euro-GBP

Esta serie es la que controla el cambio del Euro a la Esterlina (GBP). Utilizando el mismo proceso que en la serie anterior:

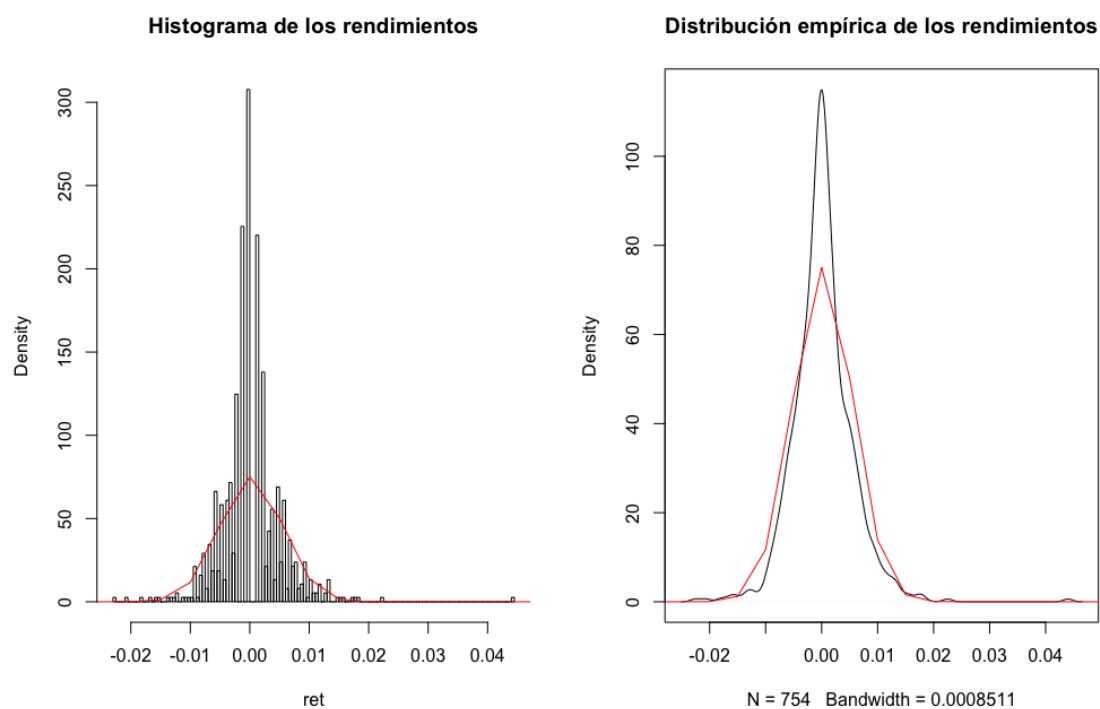


Se aprecia, igual que en la serie anterior, que la transformación tiene el aspecto adecuado poder ser modelada por procesos ARCH.

Si observamos sus funciones de autocorrelación y autocorrelación parcial:

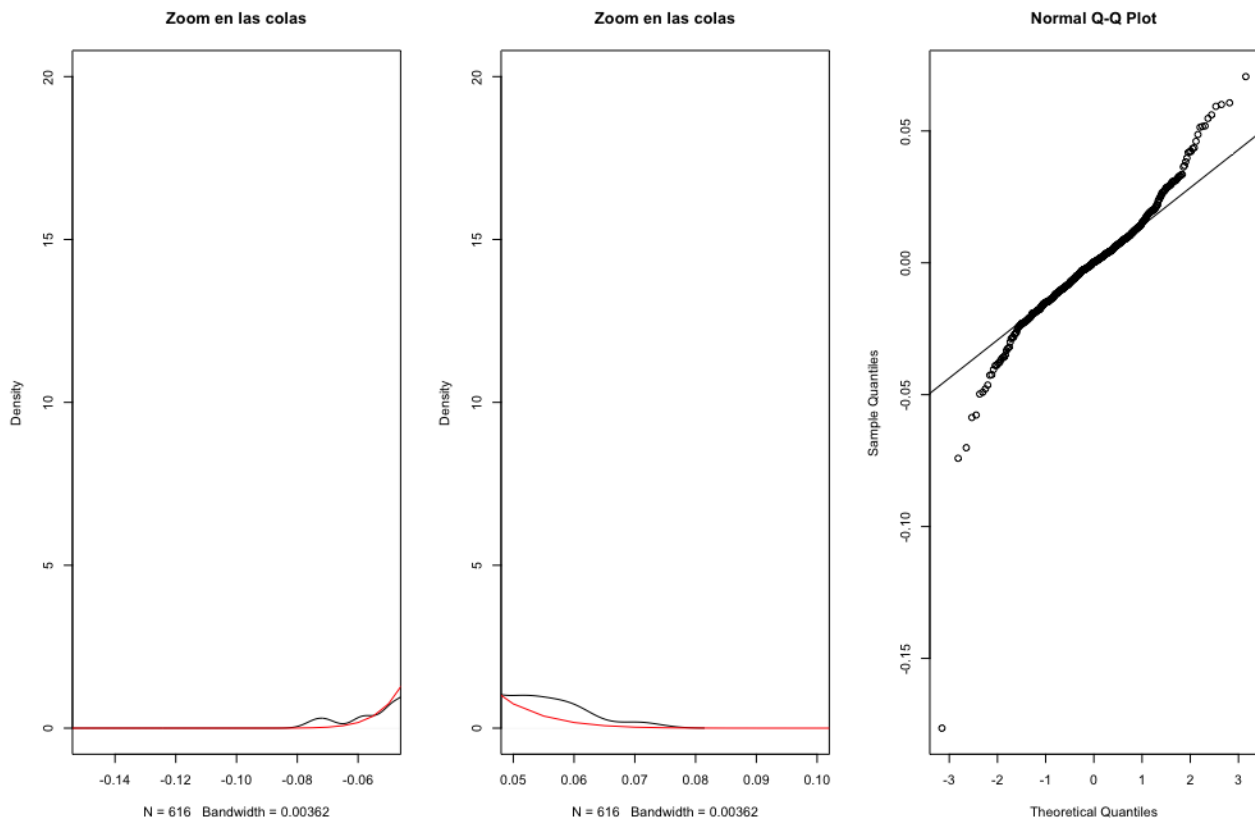


La correlación de orden 1 es inexistente, mientras que se observa una correlación cuadrática algo más destacada que en la serie anterior, aunque tampoco muy notable. Mirando el histograma y la distribución empíricas:



Vemos que la distribución empírica es apuntada y las colas son mas gruesas que las de la distribución normal. Calculando el coeficiente de curtosis, ampliando las colas y con la función QQ-plot:

```
> kurtosis(ret)
V1
14.80504
```



Se aprecia lo ya comentado, con una curtosis de 10.33878 es mucho más elevada que la de la serie anterior, y por lo tanto superior a la de la distribución normal, las colas son mas anchas y tiene un menor ajuste a la distribución normal. Se concluye que la serie se puede modelar mediante los procesos vistos en el trabajo.

Si hacemos el test Ljung-Box:

Box-Pierce test

```
data: ts^2
X-squared = 9.457, df = 1, p-value = 0.002103
```

En este caso el test sale significativo, rechazando la hipótesis nula y por lo tanto determinando que los datos no se distribuyen independientemente.

Ahora ajustaremos la serie con los modelos vistos, de los cuales el que mejor se adapta es el GARCH(1,1):

Error Analysis:

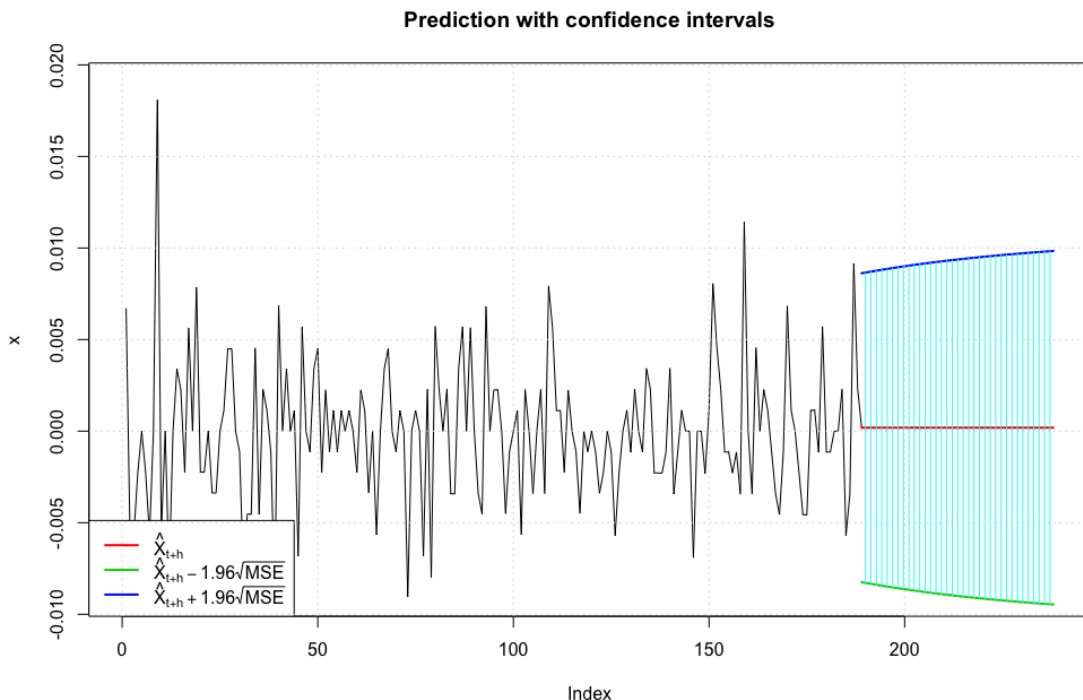
	Estimate	Std. Error	t value	Pr(> t)
mu	1.865e-04	1.761e-04	1.059	0.2894
omega	4.780e-07	5.679e-07	0.842	0.3999
alpha1	4.756e-02	2.818e-02	1.688	0.0914 .
beta1	9.358e-01	4.549e-02	20.574	<2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Vemos que la parte autorregresiva no aparece significativa al 5 %, aunque si al 10 %, por lo que la añadiremos al modelo. El modelo quedaría de la forma:

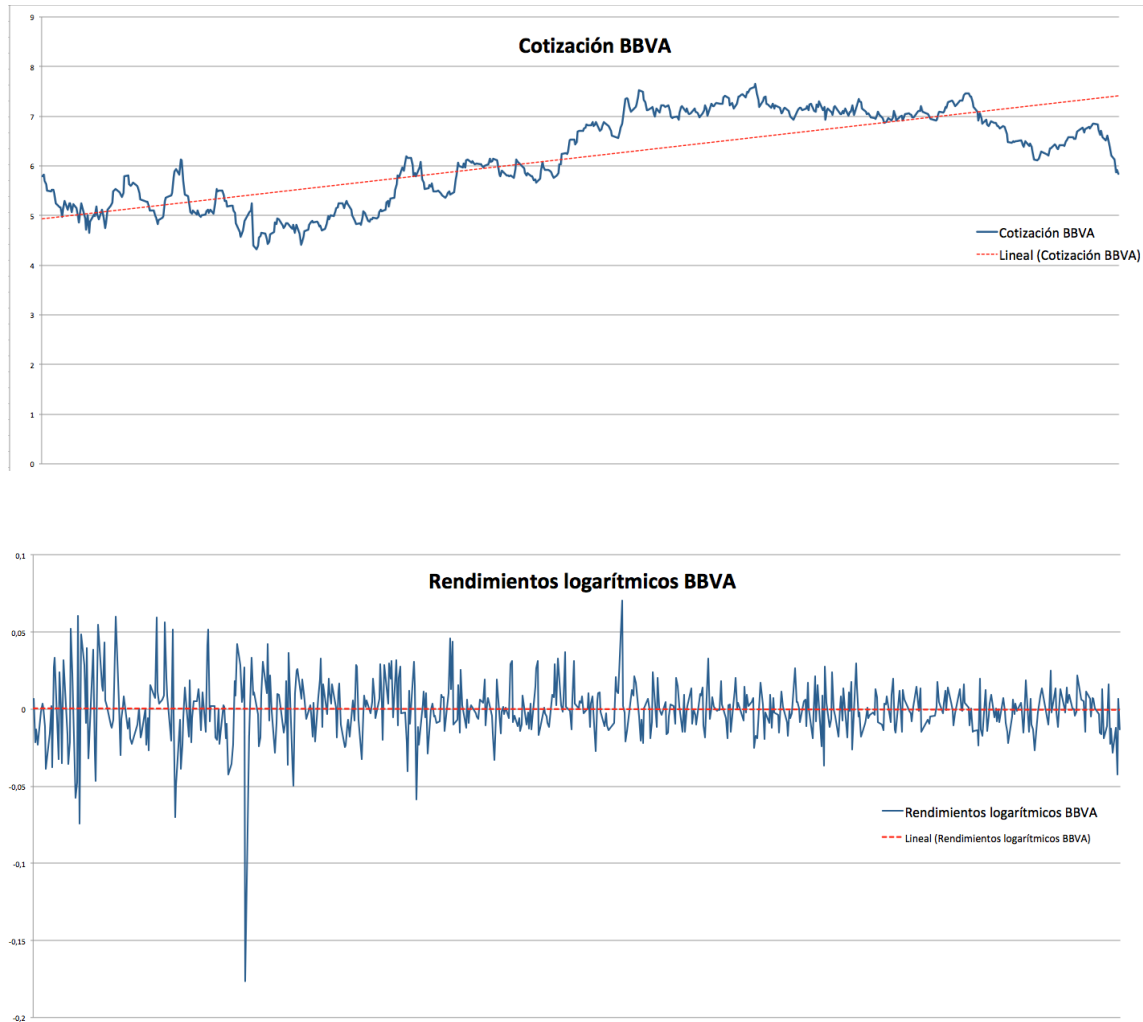
$$\sigma_t^2 = 0,04756R_{t-1}^2 + 0,9358\sigma_{t-1}^2, \quad R_t = \sigma_t Z_t \quad \text{con} \quad Z_t \sim N(0,1)$$

Y la gráfica, añadiéndole las 50 predicciones futuras, quedaría como:



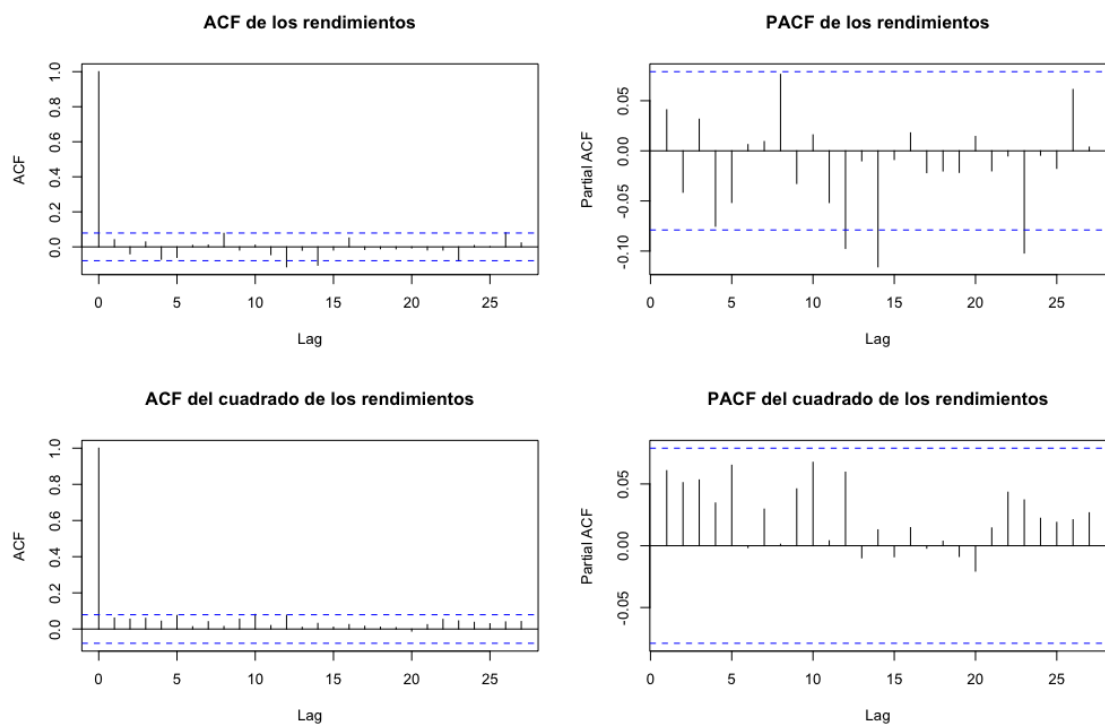
6.3. Cotización en la bolsa de BBVA

Esta serie financiera corresponde a lo que cotizan las acciones del banco BBVA. Las dos gráficas, sin y con la transformación a rendimientos logarítmicos son:

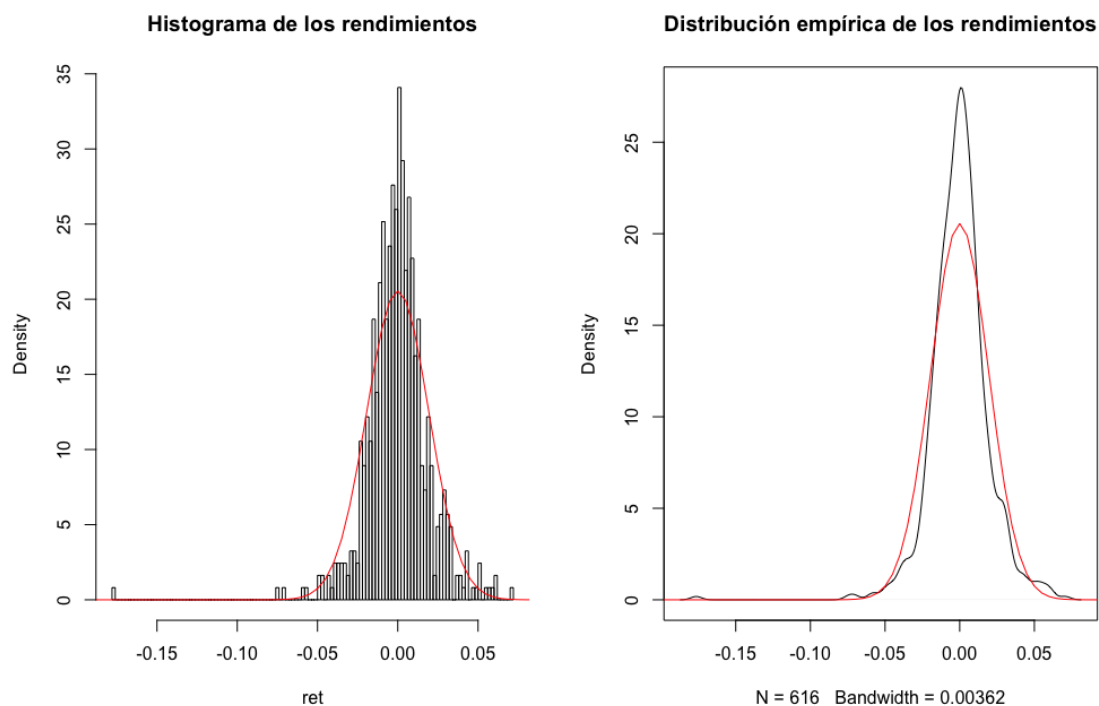


Se aprecia, igual que en la series anteriores, que la transformación tiene el aspecto adecuado poder ser modelada por procesos ARCH.

Si observamos sus funciones de autocorrelación y autocorrelación parcial:

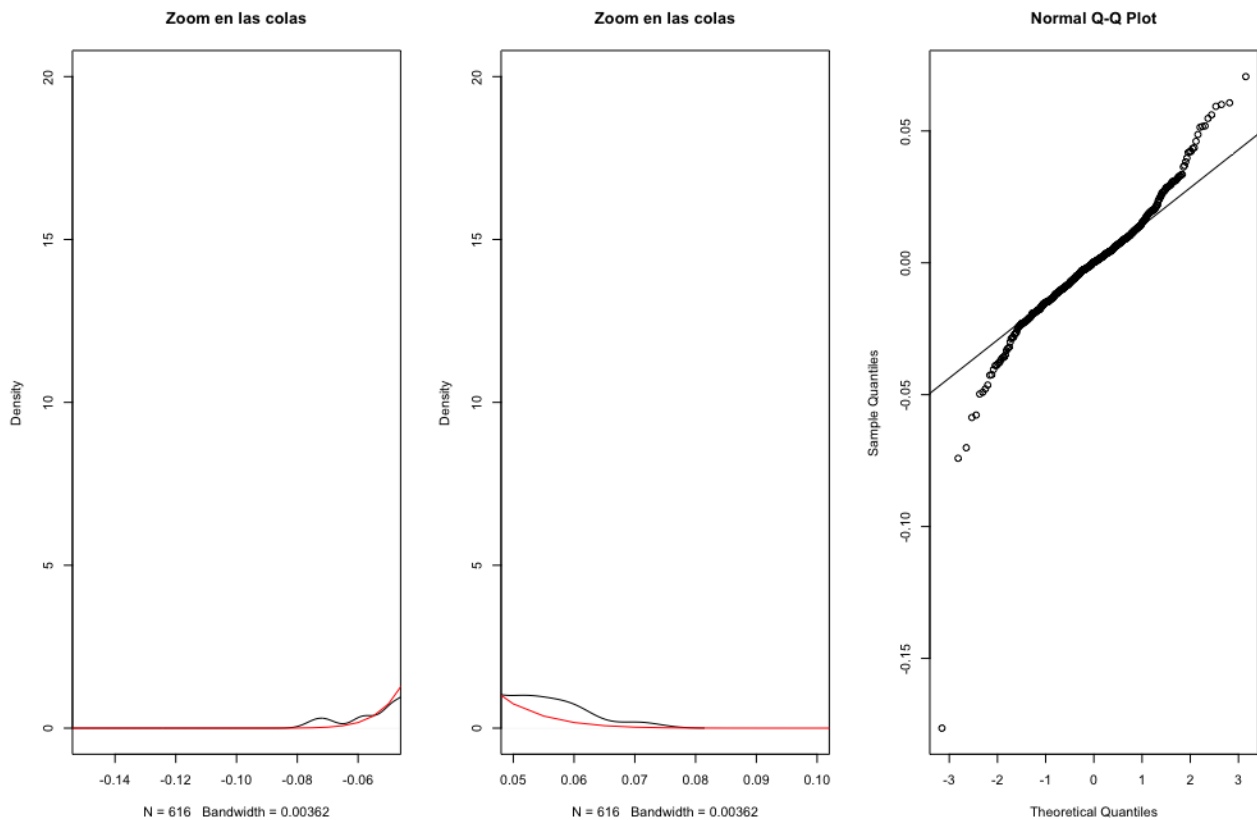


En esta serie, no se aprecia correlación ni de primer ni de segundo orden, por lo que no nos diría gran cosa. Mirando el histograma y la distribución empírica:



Vemos que la distribución empírica es mucho apuntada que la de la serie anterior, y también se aprecian unas colas mas gruesas. Calculando el coeficiente de curtosis, ampliando las colas y con la función QQ-plot:

```
> kurtosis(ret)
V1
14.80504
```



Nos aparece la curtosis más elevada de las vistas hasta ahora, con un valor de 14.80504, y las demás conclusiones serían las mismas que en la serie anterior, colas gruesas y poco ajuste a la distribución normal.

Si hacemos el test Ljung-Box:

Box-Pierce test

```
data: ts^2
X-squared = 2.2712, df = 1, p-value = 0.1318
```

En este caso el test no sale significativo, aceptando la hipótesis nula y por lo tanto determinando que los datos se distribuyen independientemente.

Ahora, pese a no ser significativo el test anterior, ajustaremos la serie con los modelos vistos, de los cuales el que mejor se adapta es nuevamente el GARCH(1,1):

Error Analysis:

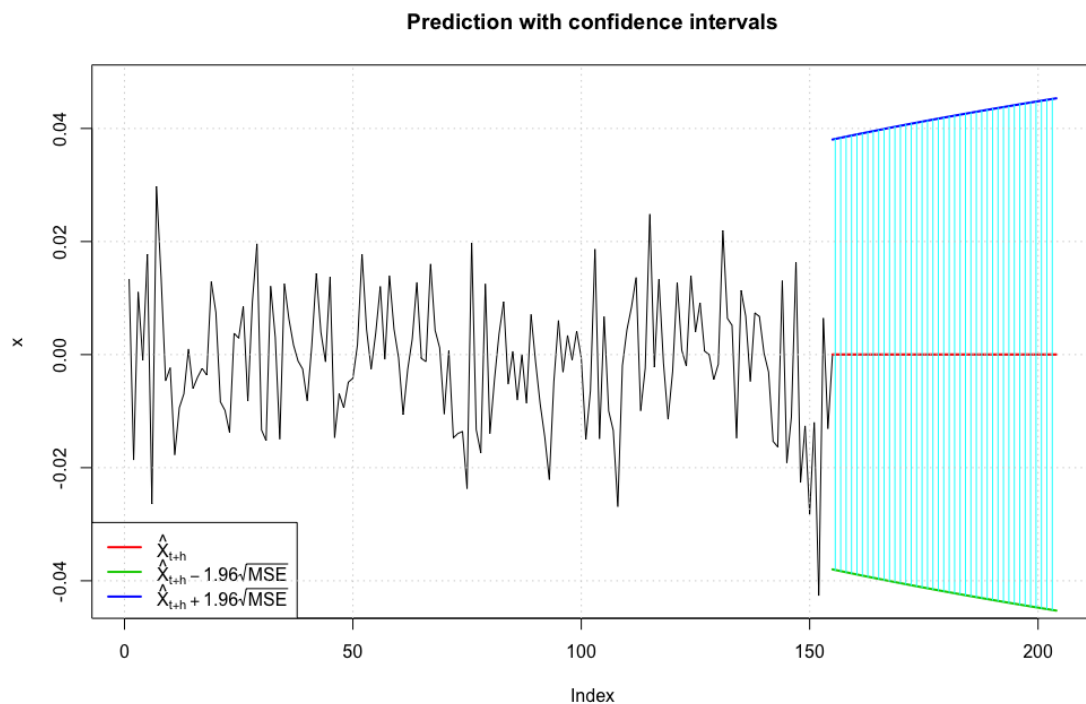
	Estimate	Std. Error	t value	Pr(> t)
mu	2.817e-05	6.097e-04	0.046	0.9631
omega	4.168e-06	4.051e-06	1.029	0.3036
alpha1	9.988e-02	4.925e-02	2.028	0.0425 *
beta1	8.980e-01	4.981e-02	18.028	<2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

El modelo quedaría de la forma:

$$\sigma_t^2 = 0,09988R_{t-1}^2 + 0,898\sigma_{t-1}^2, \quad R_t = \sigma_t Z_t \quad \text{con} \quad Z_t \sim N(0, 1)$$

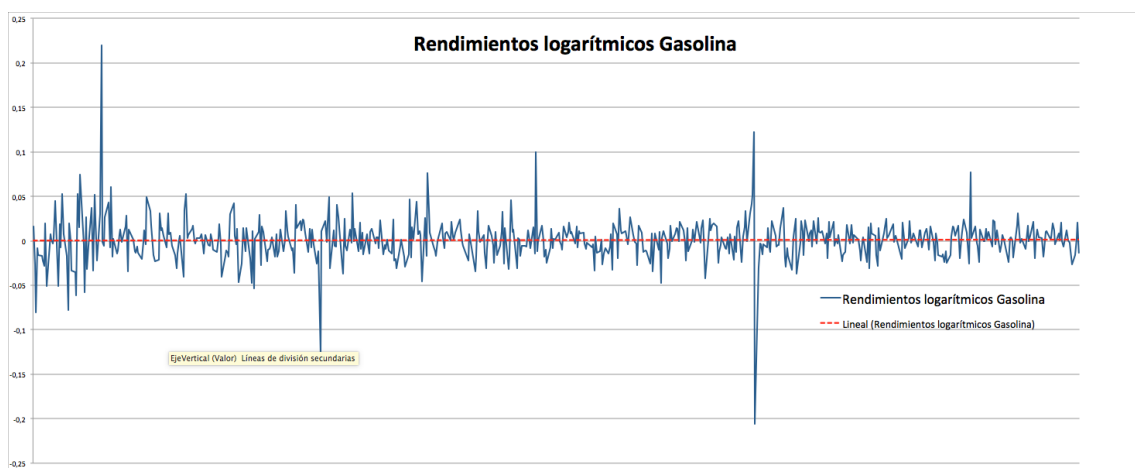
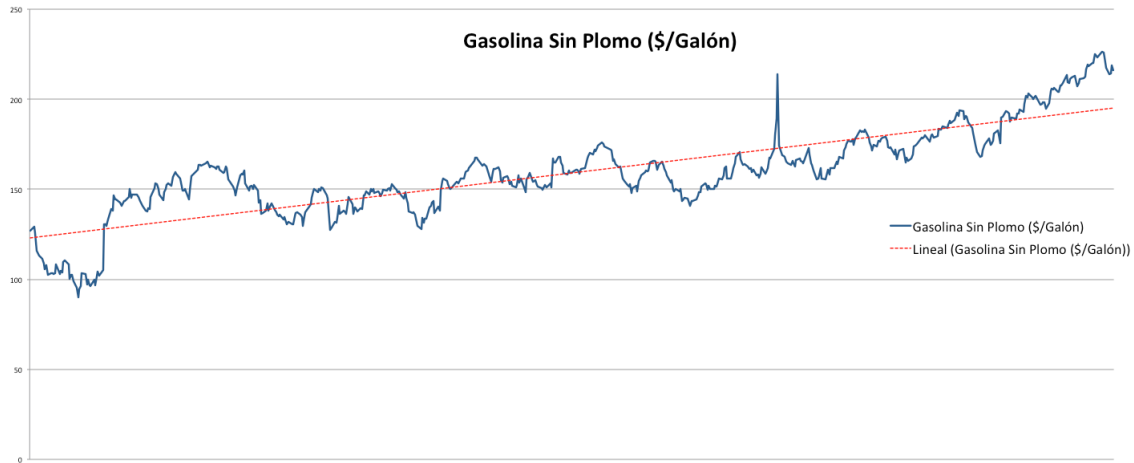
Y la gráfica, añadiéndole las 50 predicciones futuras, quedaría como:



Vemos que en esta predicción la volatilidad va en aumento, y esta lo hace de modo lineal.

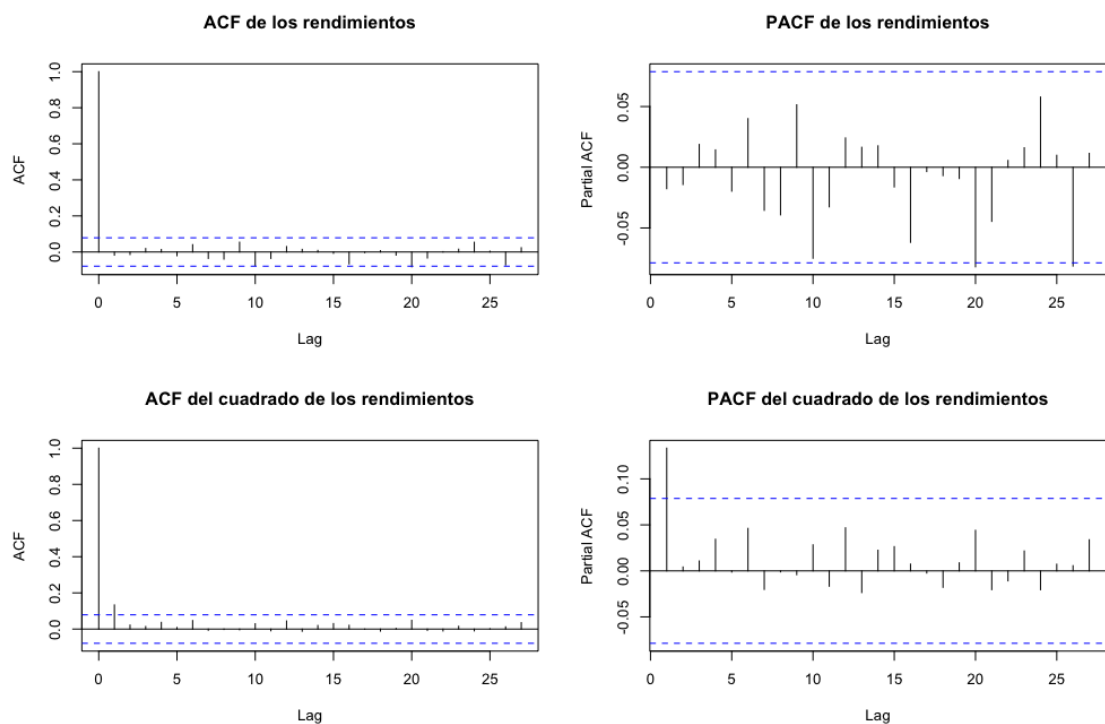
6.4. Gasolina Sin Plomo (Dólar/Galón)

Esta serie financiera corresponde al precio de la gasolina sin plomo, medido en dólares/Galón. Las dos gráficas, sin y con la transformación a rendimientos logarítmicos son:

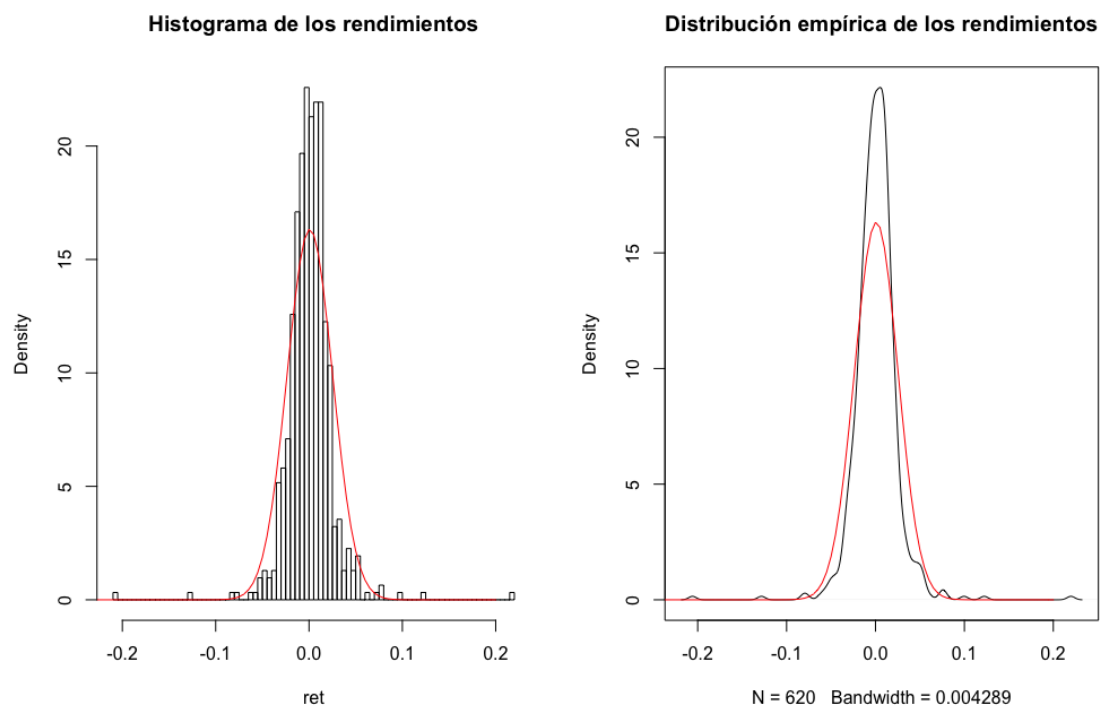


Se aprecia, igual que en la series anteriores, que la transformación tiene el aspecto adecuado poder ser modelada por procesos ARCH.

Si observamos sus funciones de autocorrelación y autocorrelación parcial:

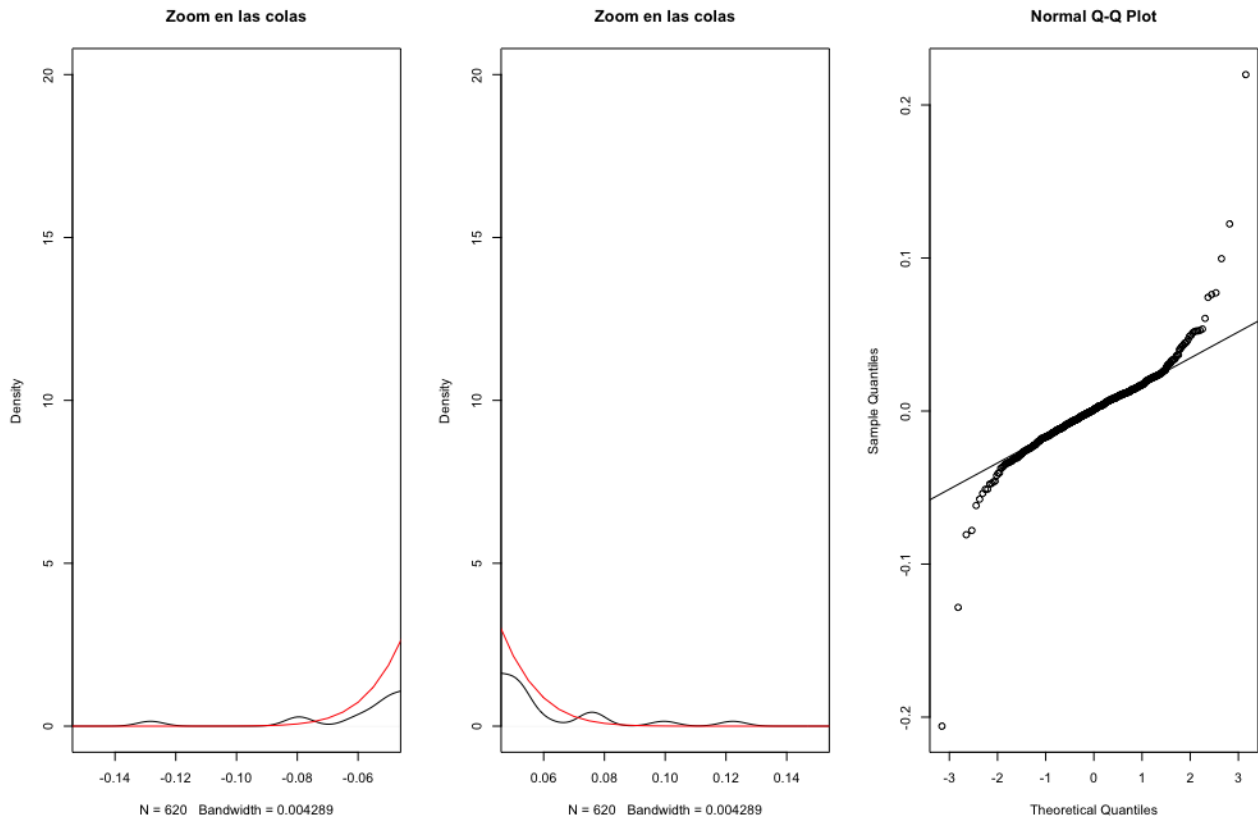


Se aprecia que la correlación de primer orden es nula, mientras que la cuadrática presenta signos de ser diferente de cero. Ahora, si observamos el histograma y la distribución emírica:



Vemos que la distribución empírica es apuntada y las colas gruesas. Calculando el coeficiente de curtosis, ampliando las colas y con la función QQ-plot:

```
> kurtosis(ret)
V1
23.44087
```



Nos aparece una curtosis mucho más elevada de las vistas hasta ahora, con un valor de 23.44087, las colas son muy gruesas y no se ajusta a la distribución normal.

Si hacemos el test Ljung-Box:

Box-Pierce test

```
data: ts^2
X-squared = 10.904, df = 1, p-value = 0.0009594
```

Es la serie en la que nos sale más significativo el test, por lo tanto rechazamos la hipótesis nula y determinamos que los datos no se distribuyen de forma independiente.

Ahora ajustaremos la serie con los modelos vistos, de los cuales el que mejor se adapta es nuevamente el GARCH(1,1):

Error Analysis:

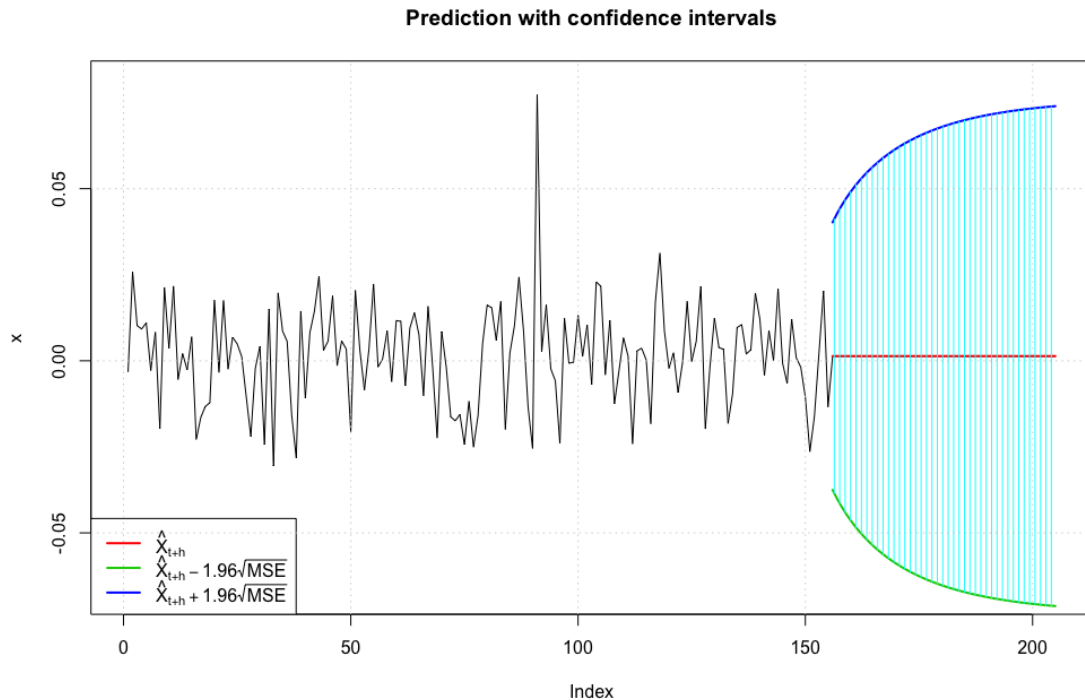
	Estimate	Std. Error	t value	Pr(> t)
mu	1.352e-03	7.878e-04	1.716	0.086240 .
omega	7.457e-05	2.204e-05	3.384	0.000716 ***
alpha1	3.692e-01	8.406e-02	4.392	1.12e-05 ***
beta1	5.794e-01	7.455e-02	7.773	7.77e-15 ***

 Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Es la única vez que nos ha aparecido el parámetro ω significativo. El modelo quedaría de la forma:

$$\sigma_t^2 = 0,00007457 + 0,3692R_{t-1}^2 + 0,5794\sigma_{t-1}^2, \quad R_t = \sigma_t Z_t \quad \text{con} \quad Z_t \sim N(0, 1)$$

Y la gráfica, añadiéndole las 50 predicciones futuras, quedaría como:



Vemos que en esta predicción la volatilidad va en aumento, pero a diferencia del anterior no lo hace de forma lineal, sino que lo hace parabólicamente.

7. Conclusiones

En general me ha parecido muy interesante hacer este trabajo, ya que he aprendido muchas cosas sobre los modelos con heterocedasticidad condicional autorregresiva.

En la parte teórica, aunque pensaba que no tendría ningún problema debido a que en la asignatura de Análisis de Series Temporales no los tuve, me he encontrado con algunas dificultades sobretodo en la parte del final. También el hecho de que los libros estén redactados en inglés ha sido un obstáculo, ya que no domino la lengua. Pese a esto, me ha parecido interesante aprender las propiedades de estos modelos y entender sus demostraciones.

Por lo que a la parte práctica respecta, ha sido entretenido buscar las diferentes series y visualizar las características que los modelos describían de primera mano, con ejemplos reales de series financieras. Además, me ha servido para aprender aun más a utilizar el programa **R** y saber usar nuevos paquetes de este.

En general, me ha gustado hacer este trabajo y esto me ha hecho interesarme aun más por estos modelos, por lo que seguiré estudiándolos pese a finalizar este trabajo.

Referencias

- [1] Franke, J.; Härdle, W. K.; Hafner, C. M.: Statistics of Financial Markets, *Springer*, January 2015.
- [2] Gouriéroux, Ch.: ARCH Models and Financial Applications, *Springer*.
- [3] Vives, J.: Pdf del Máster de Big Data, *Sin Publicar*.
- [4] Arratia, A: Computacional Finance, *Atlantis Press*.
- [5] Shumway, R. H.; Stoffer, D.S.: Time Series Analysis and Its Applications, *Springer*.
- [6] Berlinger, E; Illés, F.; Badics, M.; Banai, A; et al.: Mastering R for Quantitative Finance, *PACKT Publishing*.
- [7] Daróczi, G; Puhle, M.; Berlinguer, E.; et al.: Introduction to R for Quantitative Finance, *PACKT Publishing*.